

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号
特開2002-123837
(P2002-123837A)

(43)公開日 平成14年4月26日(2002.4.26)

(51) Int.Cl. ⁷	識別記号	F I	マークコード(参考)
G 0 6 T	15/70	G 0 6 T	15/70
1/00	3 4 0	1/00	3 4 0 A
7/20	3 0 0	7/20	3 0 0 B
			5 L 0 9 6

審査請求 未請求 請求項の数51 O.L (全 23 頁)

(21)出願番号 特願2001-261558(P2001-261558)
(22)出願日 平成13年8月30日(2001.8.30)
(31)優先権主張番号 09/651880
(32)優先日 平成12年8月30日(2000.8.30)
(33)優先権主張国 米国(U.S.)

(71)出願人 391055933
マイクロソフト コーポレイション
MICROSOFT CORPORATION
ON
アメリカ合衆国 ワシントン州 98052-
6399 レッドモンド ワン マイクロソフ
ト ウェイ (番地なし)
(72)発明者 スティーブン マーシュナー
アメリカ合衆国 98033 ワシントン州
カーランド 21 プレイス 11
(74)代理人 100077481
弁理士 谷 義一 (外2名)

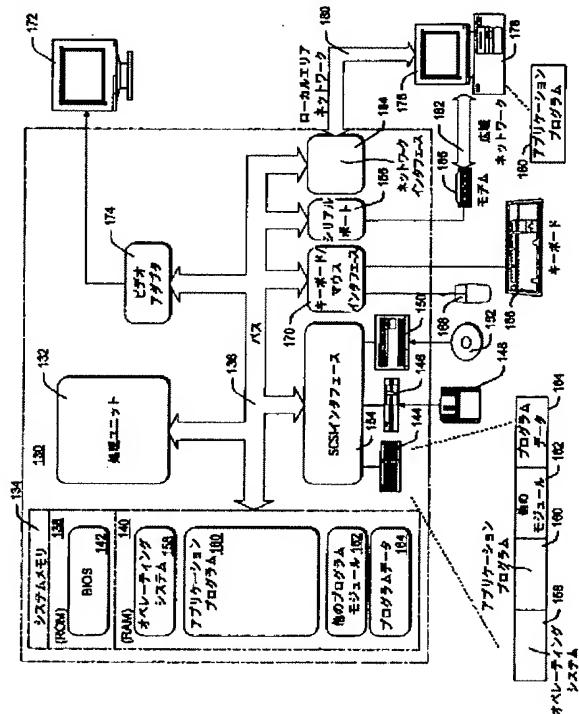
最終頁に統ぐ

(54) 【発明の名称】 頭の特徴をアニメーション化する方法およびシステムならびに表情変換のための方法およびシステム

(57) 【要約】

【課題】 顔の特徴を正確にアニメーション化し、顔の表情を変換するための方法およびシステムを提供する。

【解決手段】 コードブックは、第1人物の顔表情集合を定義しているデータを収めている。第2人物からの顔トレーニング表情集合およびコードブックからの対応する表情は変換関数を導き出すために使用され、この変換関数はコードブックの表情のすべてに適用されるようになっている。このようにして、第1人物からの表情は、第2人物の表情にリアル感をもって変換すること、およびその逆に変換することを可能にしている。多数の異なる顔を適合させる操作の基礎として使用される、单一共通汎用顔モデルが用意されている。この单一共通汎用顔モデルとユーザが定義したある種の制約を使用することにより、異なる顔間の対応付けを可能にするメカニズムが得らる。これらの変換処理は処理ユニット132で実行される。



(2)

特開2002-123837

1

【特許請求の範囲】

【請求項 1】 第1人物の第1顔表情集合を定義しているデータを収めているコードブックを定義し、
第1人物とは異なる第2人物のトレーニング表情集合を与える第2顔表情集合を定義しているデータを用意し、
前記トレーニング表情集合と、前記第1表情集合からの対応する表情とから変換関数を導き出し、
前記変換関数を該第1表情集合に適用して合成表情集合を得ることを特徴とする顔表情変換方法。

【請求項 2】 請求項1に記載の方法において、前記トレーニング表情集合は、コードブックよりも少ない表情を含んでいることを特徴とする方法。

【請求項 3】 請求項1に記載の方法において、前記変換関数は、第1人物および第2人物の顔のサイズと形状の差異を補償することを特徴とする方法。

【請求項 4】 請求項1に記載の方法において、前記変換関数を導き出すステップでは、一方の表情集合から他方の表情集合への線形的変換を計算することを特徴とする方法。

【請求項 5】 請求項1に記載の方法において、前記変換関数を導き出すステップでは、
各々の表情を、m個の標準サンプル位置でのx, y, z変位を含んでいる3mベクトルとして表現し、
変換される顔のn個の表情ベクトル集合が $g_{a1\dots n}$ として与えられ、ターゲット顔の、対応するベクトル集合が $g_{b1\dots n}$ として与えられているとき、次の形式の3m線形最小二乗システム、

【数1】

$$a_j \cdot g_a = g_b[j], i=1\dots n$$

を解くことによって、 g_a の各々の座標ごとに1つの、線形予測子集合 a_j を計算することを特徴とする方法。

【請求項 6】 請求項5に記載の方法において、前記計算するステップでは、各々の g_{aj} についてポイントの部分集合だけを使用することを特徴とする方法。

【請求項 7】 請求項6に記載の方法において、前記使用するステップでは、考慮の対象となっている標準サンプルポイントとエッジを共有しているポイントだけを使用することを特徴とする方法。

【請求項 8】 請求項5に記載の方法において、さらに、 a_j の解を求めるために擬似逆関数を計算するとき、特異値の拡散を制御することを特徴とする方法。

【請求項 9】 請求項8に記載の方法において、前記拡散を制御するステップは、 $\alpha \sigma_1$ 未満のすべての特異値をゼロとなし、ここで、 σ_1 は、マトリックスの最大特異値であることを特徴とする方法。

【請求項 10】 請求項1に記載の方法において、前記第2顔表情集合を定義しているデータを用意するステップでは、

第2人物の顔を照明で照射し、

2

顔の構造を記述している構造データと、照射からの顔の反射率特性を記述している反射率データとを同時に取り込む、

ことを特徴とする方法。

【請求項 11】 請求項10に記載の方法において、前記照射するステップでは、
複数の光源を使用し、それらの光源の1つは第2人物の顔上にパターンを投影し、そのパターンから構造データを確認可能となし、

10 前記光源の少なくとも1つは赤外線光源であり、
該光源の少なくとも1つは偏光されており、
前記取り込むステップでは、鏡面反射光を抑止して拡散成分反射データを取り込むポラライザを備えたカメラを使用することを特徴とする方法。

【請求項 12】 請求項1に記載の方法において、第2顔表情集合を定義しているデータを用意するステップでは、
顔の鏡面反射抑止反射率特性を確認できるように選択された第1偏光光源で第2人物の顔を照射し、

20 第1偏光光源で顔を照射している間に、その照射と同時に、顔上にパターンを投影する第2構造化光源で第2人物の顔を照射し、
同時照射から鏡面反射抑止反射データと構造データの両方を取り込むことを特徴とする方法。

【請求項 13】 請求項12に記載の方法において、前記第1偏光光源および前記第2構造化光源は異なる周波数の光を出力することを特徴とする方法。

【請求項 14】 請求項12に記載の方法において、前記第1偏光光源および前記第2構造化光源は赤外線光を出力することを特徴とする方法。

【請求項 15】 請求項12に記載の方法において、さらに、取り込まれたデータを処理して、(a) 顔の次元様相を記述しているデータと、(b) 顔の拡散反射特性を記述しているデータの両方を得るようにしたことを特徴とする方法。

【請求項 16】 請求項1に記載の方法において、前記第2顔表情集合を定義しているデータを用意するステップでは、

複数の異なる光源で第2人物の顔を照射し、
40 前記照射からレンジマップデータを測定し、
該照射から画像データを測定し、
前記レンジマップデータから3次元面を導き出し、
前記3次元面に対する面法線を計算し、
前記面法線と前記画像データを処理してアルベドマップを導き出すことを特徴とする方法。

【請求項 17】 請求項16に記載の方法において、複数の光源の少なくとも1つは偏光されていることを特徴とする方法。

【請求項 18】 請求項16に記載の方法において、複数の光源のすべては偏光されていることを特徴とする方

50

(3)

特開2002-123837

3

法。

【請求項 19】 コンピュータによって実行されたとき、ある人物からのトレーニング表情集合と、別の人物のコードブックからの、対応する表情とを操作し、前記トレーニング表情集合と、これらに対応する表情とから線形的変換を計算するステップ、コードブックからの複数の表情に変換関数を適用して合成表情集合を得るステップと、をコンピュータに実行させるためのコンピュータ読取可能命令を格納したことを特徴とする 1つまたは 2つ以上のコンピュータ読取可能媒体。

【請求項 20】 請求項 19 に記載のコンピュータ読取可能媒体において、前記コンピュータ読取可能命令は、コンピュータに合成表情集合を使用させて、一方の人物からの表情を他方の人物の表情に変換させることを特徴とするコンピュータ読取可能媒体。

【請求項 21】 請求項 20 に記載のコンピュータ読取可能媒体において、前記コンピュータ読取可能命令は、一方の人物からの表情であって、コードブックの表情を構成している表情とは異なっている表情を、コンピュータに変換させることを特徴とするコンピュータ読取可能媒体。

【請求項 22】 請求項 20 に記載のコンピュータ読取可能媒体において、前記コンピュータ読取可能命令は、合成表情の少なくとも 1 つのインデックスを、表情を再構築できるレシーバに送信することによって、コンピュータに表情を変換させることを特徴とするコンピュータ読取可能媒体。

【請求項 23】 請求項 20 に記載のコンピュータ読取可能媒体において、前記コンピュータ読取可能命令は、コンピュータに顔の表情を変換させることを特徴とするコンピュータ読取可能媒体。

【請求項 24】 コンピュータ読取可能媒体上に具現化されたコードブックであって、第 1 人物の第 1 顔表情集合を定義しているデータを收めているコードブックと、コンピュータ読取可能媒体上に具現化されたデータであって、該データは第 2 顔表情集合を定義しており、該第 2 顔表情集合は、前記第 1 人物とは異なる第 2 人物のトレーニング表情集合となっているデータと、

前記トレーニング表情集合と、前記第 1 表情集合からの対応する表情とから変換関数を導き出す変換プロセッサとを備えたことを特徴とする顔表情変換システム。

【請求項 25】 請求項 24 に記載の顔表情変換システムにおいて、前記変換プロセッサは線形変換プロセッサを有することを特徴とする顔表情変換システム。

【請求項 26】 請求項 24 に記載の顔表情変換システムにおいて、さらに、コンピュータ読取可能媒体上に具現化された合成表情集合を含み、該合成表情集合は、変換関数をコードブック表情に適用することによって導き

(3)

4

出されることを特徴とする顔表情変換システム。

【請求項 27】 請求項 24 に記載の顔表情変換システムにおいて、前記変換関数は、第 1 人物および第 2 人物の顔のサイズと形状の差異を補償することを特徴とする顔表情変換システム。

【請求項 28】 請求項 24 に記載の顔表情変換システムにおいて、前記変換プロセッサは、各々の表情を、m 個の標準サンプル位置での x, y, z 変位を含んでいる 3m ベクトルとして表現し、

10 変換される顔の n 個の表情ベクトル集合が $g_{a1\dots n}$ として、ターゲット顔の、対応するベクトル集合が $g_{b1\dots n}$ として与えられているとき、次の形式の 3m 線形最小二乗システム、

【数 2】

$$a_j \cdot g_a = g_b[j], i = 1..n$$

を解くことによって、 g_a の各々の座標ごとに 1 つの、線形予測子集合 a_j を計算することによって変換関数を導き出すことを特徴とする顔表情変換システム。

20 **【請求項 29】** 被写体の顔を照射するために同時に複数の異なる光源を提供する顔照明システムと、前記顔照明システムによって照射されたとき、被写体の顔から構造データと反射率データの両方を取り込むデータ取り込みシステムと、第 1 合成表情コードブックであって、該第 1 合成表情コードブックは、被写体から得られたトレーニング表情集合を受信し、前記トレーニング表情集合と、対応する未合成コードブック表情とを使用して変換関数を計算し、前記変換関数を第 1 合成情報コードブック内の表情のすべてに適用することによって合成される

30 第 1 合成情報コードブックとを有するトランスマッタと、

該トランスマッタと通信可能にリンクされたレシーバであって、顔画像を再構築するための再構築モジュールと、前記第 1 合成情報コードブックと同じ合成表情を收めている第 2 合成情報コードブックとを有するレシーバとを備え、前記トランスマッタは、被写体の追加表情を取り込み、対応するまたはほぼ合致する表情を探すために第 1 コードブックをサーチし、対応するまたはほぼ合致するコードブック表情のインデックスを、再構築モジ

40 ュールによる顔画像再構築のために前記レシーバに送信することを特徴とする顔表情変換システム。

【請求項 30】 請求項 29 に記載の顔表情変換システムにおいて、前記照明システムは、少なくとも 1 つの偏光光源を含むことを特徴とする顔表情変換システム。

【請求項 31】 請求項 29 に記載の顔表情変換システムにおいて、前記照明システムは、複数の偏光光源を含むことを特徴とする顔表情変換システム。

【請求項 32】 請求項 29 に記載の顔表情変換システムにおいて、前記照明システムは、被写体の顔の上にパターンを投影するように構成されたパターン化光源を含

50

(4)

特開2002-123837

5

むことを特徴とする顔表情変換システム。

【請求項33】 請求項29に記載の顔表情変換システムにおいて、前記照明システムは、被写体の顔の上にパターンを投影する赤外線光源を含むことを特徴とする顔表情変換システム。

【請求項34】 請求項29に記載の顔表情変換システムにおいて、前記複数の異なる光源は、すべてが赤外線光源であることを特徴とする顔表情変換システム。

【請求項35】 顔の特徴をアニメーション化する方法であって、

複数の異なる顔のジオメトリに近似している細分面を定義し、

複数の顔の各々に同じ細分面を適合させることを特徴とする顔の特徴をアニメーション化する方法。

【請求項36】 請求項35に記載の方法において、前記定義するステップでは、粗メッシュ構造をもつ細分面を定義することを特徴とする方法。

【請求項37】 請求項36に記載の方法において、粗メッシュ構造は三角メッシュを有することを特徴とする方法。

【請求項38】 請求項35に記載の方法において、前記適合させるステップでは、細分面の頂点位置にわたって連続最適化オペレーションを実行することを特徴とする方法。

【請求項39】 請求項35に記載の方法において、前記適合させるステップでは、細分面を定義しているメッシュの接続性を変更することなく、細分面を顔に適合させることを特徴とする方法。

【請求項40】 請求項35に記載の方法において、前記適合させるステップでは、細分面を定義しているメッシュに関連する平滑化汎関数を最小限にすることを特徴とする方法。

【請求項41】 請求項35に記載の方法において、前記適合させるステップでは、細分面を定義しているメッシュに関連する1つまたは2つ以上の制約を選択し、前記制約を、面上の対応するポイントに直接に適合させることを特徴とする方法。

【請求項42】 請求項41に記載の方法において、前記制約は、目、鼻および口の1つに関連付けられていることを特徴とする方法。

【請求項43】 請求項35に記載の方法において、前記適合させるステップでは、距離、滑らかさ、および制約を表す項を含む汎関数を最小限にすることを特徴とする方法。

【請求項44】 請求項35に記載の方法において、前記適合させるステップでは、一連の線形最小二乗問題の解を求めることを特徴とする方法。

【請求項45】 1つまたは2つ以上のコンピュータによって実行されたとき、当該1つまたは2つ以上のコンピュータに請求項35に記載の方法を実現させるコンピ

(4)

特開2002-123837

6

ュータ読取可能命令を格納したことを特徴とする1つまたは2つ以上のコンピュータ読取可能媒体。

【請求項46】 顔の特徴をアニメーション化する方法であって、

複数の異なる顔のジオメトリに近似する細分面を定義し、

複数の前記顔の各々に同じ細分面を適合させて、顔と顔とを対応付け、

10 顔と顔との対応付けを使用して一方の顔の表情を他方の顔の表情に変換することを特徴とする顔の特徴をアニメーション化する方法。

【請求項47】 顔の特徴をアニメーション化する方法であって、

複数の異なる顔について3次元データを測定して、対応する顔モデルを取得し、

各々の対応する顔モデルにマッピングするために使用される汎用顔モデルを1つだけ定義し、

対応する各々の顔モデル上の対応するポイントに直接にマッピングされる、前記汎用顔モデル上の複数のポイントを選択し、

対応する顔モデルの各々に該汎用顔モデルを適合させ、当該適合させるステップでは、選択されたポイントの各々を、対応する各々の顔モデル上の、対応するポイントに直接にマッピングすることを特徴とする顔の特徴をアニメーション化する方法。

【請求項48】 請求項47に記載の方法において、前記定義するステップでは、ベースメッシュ構造から細分面を定義し、前記細分面は複数の頂点を含み、顔モデルのジオメトリを近似化しており、前記適合させるステップでは、細分面の頂点の位置だけを操作することを特徴とする方法。

【請求項49】 請求項47に記載の方法において、前記適合させるステップでは、細分面を定義しているベースメッシュを操作することを特徴とする方法。

【請求項50】 請求項47に記載の方法において、前記適合させるステップでは、細分面を定義しているベースメッシュを、ベースメッシュの接続性を変更することなく操作することを特徴とする方法。

【請求項51】 請求項47に記載の方法において、前記測定するステップでは、レーザレンジスキャンを使用して3次元データを測定することを特徴とする方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、顔をリアルにアニメーション化するためのモデリングとレンダリングの方法およびシステムに関する。具体的には、本発明は顔画像処理のための方法およびシステムに関する。

【0002】

【従来の技術】コンピュータグラフィックスの分野では、さまざまなオブジェクトをレンダリングして、その

オブジェクトがユーザのためにコンピュータディスプレイから表示される。例えば、コンピュータゲームでは、コンピュータオブジェクトを生成し、レンダリングしてコンピュータモニタやテレビジョンに表示するためのコンピュータグラフィックスアプリケーションを含むことが代表的である。リアル感のある画像のモデリングとレンダリングは、コンピュータグラフィックス分野に従事するものにとっては、絶え間のない努力目標となっている。コンピュータグラフィックス分野で特に努力目標となっている分野は、リアル感のある顔画像をレンダリングすることである。1つの例として、ある特定のコンピュータグラフィックスアプリケーションでは、会話に参加している個人の表示をレンダリングすることが行われている。よく見られることは、この個人の、最終的にレンダリングされた画像はコンピュータでレンダリングされた画像であり、現実の個人とはしばしば大幅に異なっている。

【0003】

【発明が解決しようとする課題】リアル感のある顔と顔表情をモデリングし、レンダリングすることは、特に2つの理由で困難な作業となっている。1つは、人間の皮膚が、現在利用可能である、種々のシェーディングモデルでモデリングするのに不適切な反射率特性をもっているからである。例えば、周知のPhongモデルは、人間の皮膚をモデリングするには不十分である。もう1つは、顔表情をレンダリングするとき、「リアルな」顔の動きとして感知されるものと若干でもずれがあると、そのずれは、無頓着な観察者でさえも、正しくないものとして感知されるからである。現存の、顔の動き取り込み(キャプチャ)システムを使用すると、かなり説得力のある顔アニメーションを作成することができるが、取り込まれた動きは、別の顔に適用されたときは、その説得力がはるかに低下し、非常に奇異になることがよく起こっている。例えば、ある人物からその顔の動きのサンプリングをとれば、個々の顔の動きをアニメーション化することは、その動きの元となった顔が同じ顔であることを考慮すれば、困難ではない。その結果、ある表情と別の表情の間で同じか、あるいは非常に類似している動きの特徴が得られることになる。しかし、この人物の顔の動きを別の人物の顔に変換することは、この2つの顔の間には固有の差異(例えば、顔のサイズと形状)があるため、説得力に欠けることがよく起っている。

【0004】従って、本発明は、人間の皮膚のテクスチャと反射率をモデリングするための改良システムと方法を提供することを課題にしてなされたものである。また、本発明は、ある人物の顔の動きを別人物の顔の動きに変換することによって、顔の動きの取り込み(キャプチャ)データを再使用するためのシステムと方法を提供することを課題にしてなされたものである。

【0005】

【課題を解決するための手段】以下に図示し、説明している実施形態では、3次元(3-D)顔つきを記述しているデータを取り込んで、顔の動きを、ある個人から別の個人にリアルに変換し、皮膚の反射率をモデリングするための、新規な手法が提案されている。

【0006】ここで説明している実施形態では、人間の被写体が用意され、被写体の顔を照射するために複数の異なる光源が利用されている。複数の光源の中の1つは、被写体の顔の上にパターンを投影する構造化光源(structured light source)である。この構造化光源により1つまたは2つ以上のカメラが、被写体の3-Dの顔つきを記述しているデータを取り込むことが可能となる。他の光源がさらにもう1つ用意されているが、これは被写体の顔を照射するために使用されるものである。この他の1つの光源により被写体の顔の種々反射率特性(reflectance properties)を確認する。また、この他方の光源は偏光フィルタ(polarizing filter)と併用されて、顔の反射率の鏡面反射成分(specular component)が除去される。すなわち、拡散反射成分(diffuse component)だけがカメラによって取り込まれる。複数の異なる光源を使用すると、顔の構造特性と反射率特性の両方を、同時に確認することができる。光源の選択を慎重に行うと、例えば、光源をナローバンドにし、整合ナローバンドフィルタ(matching narrowband filter)をカメラで使用すると、周囲照明源の影響を除去することができる。

【0007】ここで説明している照明プロセスの結果として、2つの有用なアイテムが得られる。すなわち、(1)レンジマップ(またはデプスマップ(depth map))と、(2)構造化光源のパターンがそこに投影されていない顔の画像である。3D面がレンジマップから導き出され、3D面に対する面法線(surface normal)が計算される。3D面を定義するためにレンジマップを処理するとき、その処理にオプションとしてのフィルタリングステップを含めておくと、汎用顔テンプレート(generic face template)がレンジマップと結合され、不要なノイズが拒絶されることになる。計算された面法線と顔の画像は、アルベドマップ(albedo map)を導き出すために使用される。アルベドマップは特殊タイプのテクスチャマップ(texture map)であり、そこでは、各サンプルは、面上の特定ポイント(点)における顔面の拡散反射率を記述している。従って、プロセスのこの時点で、顔の3D面(つまり、面法線)を記述している情報と、顔の反射率(つまり、アルベドマップ)を記述している情報は確認されていることになる。

【0008】ある実施形態では、照明プロセスで得られた情報またはデータは、ある人物の顔の表情を、別人物の顔の表情に変換するために使用される。この実施形態では、コードブック(code book)の考え方方が取り入れられ、使用されている。

【0009】コードブックは、別人物（人物A）の、多数の汎用表情（generic expressions）を記述しているデータを収めている。1つの目標は、コードブックの表情を取り出し、その表情を使用して別人物（人物B）の表情を変換することである。これを行うために、本発明による方法では、人物Bを使用してトレーニング表情（training expressions）の集合を作成している。トレーニング表情は、コードブックに存在する表情の集合から構成されている。トレーニング表情および各表情に対応するコードブック表情を使用すると、変換関数（transformation function）が導き出される。この変換関数は、人物Bの表情に合致しているはずの合成表情の集合（a set of synthetic expressions）を導き出すために使用される。すなわち、変換関数が導き出されると、これはコードブック内の表情の各々に適用され、コードブック表情が人物Bの表情に合致するようにされる。従って、トレーニング集合の中に存在しない可能性のある、新しい表情が、例えば、人物Bから受信されたときは、合成コードブック表情をサーチすれば、人物Bの表情に最も合致する表情を見付けることができる。

【0010】別の実施形態では、顔の表情と動きを、ある顔から別の顔に変換するために使用できる共通顔構造（common face structure）が定義されている。ここで説明している実施形態では、共通顔構造は、粗メッシュ構造（coarse mesh structure）または「ベースメッシュ（base mesh）」からなり、これは、ある人物の表情を別の人物の表情に変換するときの基礎として使用される細分面（subdivision surface）を定義している。共通ベースメッシュはすべての顔に対して使用され、2または3以上の顔が対応付けられるようにしている。従って、これによって定義された構造を使用すると、顔の動きをある人物から別人物に適応することが可能になる。この実施形態によれば、細分面を被写体の顔モデルに適応するための手法が用いられている。本発明による手法では、細分面上の、いくつかのポイントは、顔モデル上の対応するポイントに直接的にマッピングされるように定義されている。このことは、起こり得る、あらゆる異なる顔モデルについても同じである。この制約（constraint）を加えると、ベースメッシュは、異なる顔モデルを同じように適合させる、という特性をもつことになる。さらに、本発明によるアルゴリズムでは、ベースメッシュと顔モデルとが確実に対応づけられるように最小限にされた平滑化汎関数（smoothing functional）が利用されている。

【0011】別の実施形態では、被写体の顔面の反射率を計量化する反射率処理手法が用意されている。反射率を測定するために、本発明による手法では、反射率は拡散反射成分（diffuse component）と鏡面反射成分（specular component）に分離され、拡散反射成分の処理が中心になっている。

【0012】拡散反射成分を測定するために、アルベドマップ（albedo map）が最初に定義される。このアルベドマップは、カメラと、複数の異なる光源によって照射される被写体とを、最初に用意することによって定義される。光源は偏光フィルタ（polarizing filter）によってフィルタリングされる。このフィルタリングは、偏光フィルタがカメラの前に置かれた偏光フィルタと併用されて、鏡面反射を抑止するか、あるいは鏡面反射が記録されるのを防止することによって行われる。画像シーケンスは、被写体の頭を取り巻くように撮影される。各々の個別画像は、その画像に対応する個別アルベドマップが得られるように処理される。特定の被写体に対するすべてのアルベドマップは、その被写体の顔全体に対して単一のアルベドマップが得られるように結合される。

【0013】

【発明の実施の形態】以下、図面を参照して本発明の実施形態を詳細に説明する。

【0014】概要

リアル感のある顔と顔表情をレンダリングするには、皮膚の反射率と顔の動きを非常に正確なモデルで表す必要がある。以下では、ジオメトリ、動きおよび反射率に関する測定データを使用して、顔をモデリングし、アニメーション化し、レンダリングして特定人物の顔の外観と顔表情をリアルに再現するための方法および手法が説明されている。ジオメトリと双方向反射率を含んでいる完全モデルが構築されるので、どのような照明条件と表示条件においても、顔をレンダリングすることができる。以下に説明するモデリングシステムと方法では、異なる顔間が対応付けられた構造化顔モデル（structured face model）が構築され、この顔モデルは、さまざまな顔のアニメーション操作の基礎となっている。

【0015】本発明による実施形態は、以下に説明するように、顔モデリングプロセスを構成する各部分と係わりをもっている。ジオメトリの表現を構造化し、統一化することによって、そのジオメトリ表現を顔モデルの基礎とし、多くの以後の顔モデリングとレンダリング操作の基礎となるようにするために、本発明の実施形態では、従来の面適合手法（surface fitting technique）を拡張し、汎用の顔（generic face）が異なる個人の顔に合致させるようにしている。リアル感のある反射率モデルを構築するために、最初に、最近の皮膚反射率測定が使用されるが、これは公知で実用化されている。さらに、改良されたテクスチャ取り込みプロセス（texture capture process）を使用して、新規に測定された拡散テクスチャマップが付加されている。汎用メッシュ（generic mesh）をアニメーション化するために、改良された手法が使用され、高品質レンダリングに適した面形状が得られるようになっている。

【0016】例示コンピュータシステム

50 予備知識として、図1に、ここで説明している実施形態

で使用できるデスクトップコンピュータ130の例の概要を示す。分散コンピューティング環境の状況下では、図示のようなコンピュータはさまざまなもののが使用可能である。これらのコンピュータは、以下に記載されている説明に従ってグラフィックスをレンダリングし、画像（イメージ）を処理するために使用することができる。

【0017】コンピュータ130は、1つまたは2つ以上のプロセッサまたは処理ユニット132、システムメモリ134、およびバス136、すなわち、システムメモリ134を含む種々のシステムコンポーネントをプロセッサ132に結合しているバス136を含んでいる。バス136は、数種タイプのバス構造のいずれか1つまたは2つ以上で表されているが、この中には、さまざまなバスアーキテクチャを採用したメモリバスまたはメモリコントローラ、ペリフェラルバス、高速グラフィックスポート、およびプロセッサまたはローカルバスが含まれている。システムメモリ134としては、リードオンリメモリ（read only memory - ROM）138とランダムアクセスメモリ（random access memory - RAM）140がある。基本入出力システム（basic input/output system - BIOS）142は、スタートアップ時のときのように、コンピュータ内のエレメント間の情報転送を支援する基本ルーチンで構成され、ROM138に格納されている。

【0018】コンピュータ130は、さらに、ハードディスク（図示せず）との間で読み書きするハードディスクドライブ144、取り外し可能ディスク148との間で読み書きする磁気ディスクドライブ146、CD-ROMや他の光媒体のような、取り外し可能光ディスク152との間で読み書きする光ディスクドライブ150を装備している。ハードディスクドライブ144、磁気ディスクドライブ146、および光ディスクドライブ150は、SCSIインターフェース154または他の、ある種の該当ペリフェラルインターフェースを通してバス136に接続されている。これらのドライブおよびその関連コンピュータ読取可能媒体は不揮発性ストレージとして、コンピュータ読取可能命令、データ構造、プログラムモジュールおよび他のデータを、コンピュータ130のために格納している。ここで説明している例示環境では、ハードディスク、取り外し可能磁気ディスク148および取り外し可能光ディスク152が採用されているが、当業者ならば理解されるように、磁気カセット、フラッシュメモリカード、デジタルビデオディスク、ランダムアクセスメモリ（RAM）、リードオンリメモリ（ROM）などのように、コンピュータがアクセスできるデータを格納しておくことができる、他のタイプのコンピュータ読取可能媒体を、例示の動作環境で使用することも可能である。

【0019】いくつかのプログラムモジュールは、ハードディスク144、磁気ディスク148、光ディスク152、ROM138またはRAM140に格納しておくこと

が可能であり、その中には、オペレーティングシステム158、1または2以上のアプリケーションプログラム160、他のプログラムモジュール162、およびプログラムデータ164が含まれている。ユーザは、キーボード166やポインティングデバイス168などの入力デバイスを通して、コマンドや情報をコンピュータ130に入力することができる。他の入力デバイス（図示せず）として可能なものとして、マイクロホン、ジョイスティック、ゲームパッド、サテライトディッシュ、スキヤナ、1つまたは2つ以上のカメラなどがある。上記および他の入力デバイスは、バス136に結合されたインターフェース170を通して処理ユニット132に接続されている。モニタ172または他のタイプのディスプレイデバイスも、ビデオアダプタ174などのインターフェースを介してバス136に接続されている。モニタのほかに、パーソナルコンピュータは、スピーカやプリンタのような、他の周辺出力デバイス（図示せず）を装備しているのが代表的である。

【0020】コンピュータ130は、リモートコンピュータ176のような、1つまたは2つ以上のリモートコンピュータとの論理コネクションを使用するネットワーキング環境で動作するのが通常である。リモートコンピュータ176としては、別のパーソナルコンピュータ、サーバ、ルータ、ネットワークPC、ピアデバイスまたは他の共通ネットワークノードにすることが可能であり、図1にはメモリストレージデバイス178だけが示されているが、コンピュータ130に関連して上述したエレメントの多くまたはすべてを装備しているのが代表的である。図1に示す論理コネクションとしては、ローカルエリアネットワーク（local area network - LAN）180と広域ネットワーク（wide area network - WAN）182がある。この種のネットワーキング環境は、オフィス、企業内コンピュータネットワーク、インターネット、およびインターネット（the Internet）ではありふれた存在になっている。

【0021】LANネットワーキング環境で使用されるときは、コンピュータ130は、ネットワークインターフェースまたはアダプタ184を通してローカルネットワーク180に接続されている。WANネットワーキング環境で使用されるときは、コンピュータ130は、インターネットのような、広域ネットワーク182上のコミュニケーションを確立するために、モデム186またはネットワークインターフェースのような他の手段を装備しているのが代表的である。モデム186は内蔵型と外付け型があり、どちらも、シリアルポートインターフェース156を介してバス136に接続されている。ネットワーキング環境では、パーソナルコンピュータ130に関連して上述したプログラムモジュールまたはその一部は、リモートメモリストレージデバイスに置いておくことが可能である。当然に理解されるように、図示のネットワー-

(8)

13

クコネクションは例示であり、コンピュータ間の通信リンクを確立する他の手段を使用することも可能である。

【0022】一般的に、コンピュータ130のデータプロセッサは、コンピュータの種々コンピュータ読取可能媒体に、それぞれ時間を異にして格納されている命令によってプログラムされている。プログラムとオペレーティングシステムは、例えば、フロッピディスクやCD-ROM上に分散化されているのが代表的である。プログラムとオペレーティングシステムは、フロッピディスクやCD-ROMからコンピュータの二次メモリにインストールまたはロードされる。実行時には、プログラムとオペレーティングシステムは、少なくともその一部が、コンピュータの一次エレクトロニックメモリにロードされる。ここで説明している本発明によれば、マイクロプロセッサまたは他のデータプロセッサに関連して以下に説明されているステップを実現するための命令またはプログラムは、上述したタイプまたは他の種々タイプのコンピュータ読取可能記憶媒体に格納されている。また、本発明によれば、コンピュータ自体は、以下に説明されている方法および手法に従ってプログラムされている。

【0023】説明の便宜上、オペレーティングシステムのように、プログラムや他の実行可能プログラムコンポーネントは、ここでは離散的ブロックとして示されているが、当然に理解されるように、これらのプログラムとコンポーネントは、時期を異にしてコンピュータの異種ストレージコンポーネントに置かれていて、コンピュータのデータプロセッサによって実行される。

【0024】顔面の構造と特性を取り込むための例示システム

従来は、取り込みシステム(capturing system)は、顔全体の顔構造と反射率特性の両方を、相互から独立して同時に取り込むことができなかつた。例えば、構造化光(structured light)を使用して顔の構造を取り込むシステムが存在しているが、これらのシステムは、反射率(reflectance)のような、顔の特性を取り込んでいない。同様に、顔の反射率を取り込むシステムが存在しているが、これらのシステムは、顔の構造を取り込んでいない。顔の構造と反射率を、相互に独立して同時に取り込むことができると、収集したデータに別の操作を加え、さまざまな顔レンダリングとアニメーション化操作に役立てることが可能になる。以下では、例示レンダリング操作の一具体例が説明されているが、当然に理解されるように、下述しているシステムと方法の結果として得られる情報またはデータは、他のさまざまな分野で利用することが可能である。例えば、アプリケーション分野としては、セキュリティやパーソナルユーザ間のやりとり、などを目的とした顔の認識、ゲーム、ムービー(映画、動画)などのアニメーションを目的としたリアル感のある顔モデルの構築、およびユーザが自分の顔を容易に取り込んで、インターフェイブエンターテイメントや

特開2002-123837

14

ビジネスコミュニケーションで使用できるようにすること、などがある。

【0025】図2は、被写体の顔の顔構造と反射率特性を同時に取り込む上で、使用するのに適している例示システム200を示す。この例示システムは、1つまたは2つ以上のカメラの形体をしたデータ取り込みシステムを装備し、その1つがカメラ202として例示されている。カメラ202は、CCD画像センサと関連回路を装備し、CCDアレイを動作させ、CCDアレイから画像を読み取り、画像をデジタル形体に変換し、その画像をコンピュータに通知することを可能にしている。また、システムは、複数の光源またはプロジェクタの形体をした顔照明システム(facial illumination system)も装備している。複数のカメラが使用される場合には、これらのカメラは、時間的に同時取り込みを可能にするようにジェンロック(genlock)されている。図示の例では、2つの光源204、206が利用されている。光源204は、望ましくは、被写体の顔の上に投影される構造化パターン(structured pattern)を出力する。光源204は、20適當なロケーションであれば、どこにでも置くことが可能になっている。このパターンは、被写体の顔の3D形状に関する構造情報またはデータを、カメラ202で取り込むことを可能にしている。赤外線領域の光で構成されたパターンを採用すると有利であるが、適當な光源であれば、どの光源でも使用可能である。光源206は、望ましくは、顔の反射率特性の拡散反射成分をカメラ202で取り込めるようにする光を出力する。光源206は、例えば、ビームスプリット(beam splitting)手法でカメラのレンズ202aと同列に置いておくと有利であるが、適當なロケーションであれば、どこにでも置くことが可能である。この光源は、カメラのレンズを取り囲むように適応させることも可能である。また、この光源は、反射率の鏡面反射成分(specular component)が抑止または除去されるように選択されている。図示の例では、線形偏光フィルタ(linear polarizing filter)は偏光照明を得るために採用され、別の偏光フィルタは前記線形偏光フィルタに直交する向きになっていて、レンズ202aの前に置かれているため、顔からの鏡面反射がカメラによって記録されないようにしている。上述した照明システムは、異なる周波数の光源、例えば、カメラの赤と緑チャネルに対応する光源を使用してシミュレートされている。なお、これらのチャネルは、どちらも赤外線領域にすることが可能である。さらに、光源がナローバンド(例えば、780-880 nm)になるように選択すると、周囲光の影響を除去することができる。この特性が達成されるのは、カメラもナローバンドになるようにフィルタリングされるときだけである。光源からの照明はナローバンドの波長に集中されるのに対し、周囲光はブロードレンジの波長にわたって拡散されるので、光源からの光は、上記のような波長のときの周囲光よりも

(9)

15

強力になる。カメラは、光源から放出された波長だけを記録するようにフィルタリングされるので、周囲光による影響を受けることはほとんどない。その結果、カメラは、選択された光源が被写体に及ぼす影響だけを検出することになる。

【0026】複数の異なる光源を使用すると、具体的には、赤外線光源を偏光光源（これも赤外線光源にすることが可能である）と併用すると、カメラは、顔に関する構造情報またはデータ（光源204からの）と顔に関する反射率情報またはデータ（光源206からの）を、相互に独立して同時に取り込むことが可能になる。構造情報は顔の3次元様相を記述しているのに対し、反射率情報は顔の拡散反射率特性を記述している。この情報は、コンピュータ208などの、コンピュータ化画像プロセッサによって処理されると、以後の顔アニメーション操作で使用できる情報またはデータが得られることになる。以下で説明する例では、この情報は3次元データ（3Dデータ）とアルベドマップ（albedo map）を含んでいる。

【0027】図3は、ここで説明する実施形態による方法のステップを記載しているフローチャートである。ここで説明する方法によれば、顔の構造と反射率特性に関する情報またはデータは、同時に収集し、処理することが可能になっている。ステップ300で、被写体の顔は複数の異なる光源で照射される。このステップを実現する例示システムは図2に示されている。なお、図示の例では、2つの例示光源が利用されているが、当然に理解されるように、他の複数の光源を使用することも可能である。ステップ302で、ステップ300の照射からレンジマップデータ（デプスマップデータ）と画像データが測定される。すなわち、ステップ300の照射により、カメラは光反射率を検出することが可能になり、この光反射率は、レンジマップデータと、そこに構造光源パターンを含んでいない画像データ（つまり、反射率）の両方を得るために利用されることになる。レンジマップデータと画像データはコンピュータ208（図2）に渡され、そこで処理される。この時点で、ステップ304で、汎用顔テンプレート（generic face template）をレンジマップデータに適用すると、レンジマップデータに関連して生じている種々のノイズを除去することができるが、これはオプションである。汎用顔テンプレートは、レンジマップデータに含まれるノイズを除去する3Dフィルタと考えることができる。汎用顔テンプレートは、この分野の精通者に自明のものである。

【0028】ステップ306で、レンジマップデータが使用され、3D面が導き出され、あるいは計算される。適当なアルゴリズムであれば、どのアルゴリズムでも使用可能であり、これは、この分野の精通者に自明のものである。アルゴリズムの例としては、次の論文に記載されているものがある。すなわち、Turk & Levoy著「レンジ

特開2002-123837

16

画像からのジッパドポリゴンメッシュ（Zippered Polygon Meshes）」（SIGGRAPH94）、およびF. Bernardini, J. Mittleman, H. Rushmeier, C. Silva, G. Taubin共著

「面再構築のためのボールビボッティングアルゴリズム（The Ball-Pivoting Algorithm）」（Trans. Vis. Comp. Graph. 5:4 (1999)）である。ステップ308で、ステップ306の3D面に対する面法線ベクトル（「面法線」）が、公知のアルゴリズムを使用して計算される。このタスクを実行する1つの方法は、三角形に対する法線を計算し、各頂点周りの三角形法線の平均をとって、頂点法線を作り、頂点法線を各三角形の内部を横切るように補間することである。当然のことであるが、他の方法を利用することも可能である。次に、ステップ310で、ステップ308で計算された面法線とステップ302の画像データが使用されて、アルベドマップが導き出される。アルベドとは、特殊タイプのテクスチャマップであり、そこでは、各サンプルは、顔面上の特定ポイントでの顔面の拡散反射率を記述している。アルベドマップの導出は、上記に示した情報が与えられていれば、この分野の精通者には自明のものである。例示のアルゴリズムは、Marschner著「コンピュータグラフィックスの逆レンダリング（Inverse Rendering for Computer Graphics）」（PhD thesis, Cornell University, August 1998）に記載されている。

【0029】この時点で、図2に示すように、照明処理の結果として、被写体の顔の構造特徴を記述している3Dデータと、画面の拡散反射率を記述しているアルベドマップデータとが得られている。

【0030】上記照明処理は、記述されている情報を抽出するために使用できるので、抽出した情報は適当な目的のために使用することができる。特に利点のある実施形態では、抽出された情報は、被写体の表情を抽出し、認識するために利用されている。この情報は、表情変換のために使用できる。すぐあとで説明されている、本発明による実施形態では、ある人物の表情は、別人物の表情を、リアル感をもって変換するために使用することができる。

【0031】コードブックの使用による表情変換
ある表情変換実施形態では、コードブック（code book）40の考え方方が取り入れられ、以下に説明されている表情変換操作で利用されている。図4は、ある人物から取り込まれた多数の、異なる表情を収めているコードブック400の例を示す。これらの表情は、汎用表情（generic expressions）、つまり、特定の個人からではなく、汎用的個人からの表情と考えることができる。この例では、表情は、表情1から表情Nまでの範囲を有する。表情1は、例えば、「スマイル」であり、表情2は「しかめ面」であり、表情3は「怒り」の表情である。以下、同様である。コードブック400に収められている表情は、数学的には、それらのジオメトリの形で記述さ

れており、すぐ上で説明したプロセスのような、適当な方法で取り込むことが可能である。

【0032】表情変換を行うために、コードブック400内の表情の一部を使用して、変換関数(transformation function)が最初に導き出される。変換関数を導き出すために、トレーニング表情集合(training set of expressions)402という考え方を取り入れられている。表情トレーニング集合402は、その表情がコードブック400に記述されている個人とは別の個人から得られる表情の集合で構成されている。トレーニング集合402のトレーニング表情は、コードブック表情のサブセット(部分集合)になっている。つまり、トレーニング集合の中の各表情は、コードブック400の中の表情に対応している。例えば、トレーニング集合402は、3つの表情、すなわち、表情1、表情2、および表情3で構成することができる、そこでは、表情は、それぞれ「スマイル」、「しかめ面」および「怒り」になっている。変換関数の目標は、トレーニング集合の表情に関連付けられたジオメトリック変形を受け取り、そのジオメトリック変形をコードブック400の表情のすべてに適用して、コードブック表情がリアル感のある表情表現となるようにすることである。すなわち、各人物の顔は、どの任意の表情についても、ジオメトリが異なった変形の仕方をするものと考えられる。ある表情のときの、ある人物のジオメトリックな顔の変形が、表情レンダリングする目的で別人物の顔に単純に適用されると、その変形が適用された顔は、非常にゆがんだように見えることになる。このようなことが起こるのは、顔のジオメトリが異なるだけでなく、顔と顔の間のように顔の変形が異なるためである。従って、ある表情集合から別の表情集合への最良変換が得られるような変換関数が導き出されている。

【0033】図4に戻って説明すると、図4には、線形変換プロセッサ406が示されている。この変換プロセッサ406は、適当なハードウェア、ソフトウェア、ファームウェア、あるいはこれらの組み合わせで実現することが可能である。図示の例では、線形変換プロセッサ406はソフトウェアで実現されている。線形変換プロセッサはトレーニング表情集合402と、対応するコードブック表情404とを入力として受け取る。変換プロセッサはこれらの入力を処理して、変換関数408を導き出す。この変換関数408を、コードブック400内の表情のすべてに適用すると、合成表情集合(synthesized set of expressions)が得られる。この合成表情集合は、トレーニング表情集合が得られた人物の表情に関連するジオメトリック変形によって操作された、コードブックの表情を表している。

【0034】表情が同じであるときの顔の変位は、次の2つの理由で異なる人物では同じにならない。その1つは、動き取り込みサンプルポイント(この具体例で

は、この特定アルゴリズムによると、顔の動きがどのように表現できるかを示している)は、位置に誤差があると、正確に一致しないからである。もう1つは、頭の形状とサイズは人物と人物の間で変化するからである。

【0035】第1の不一致は、すべての顔の動き取り込み変位データを、汎用メッシュ上の一定位置集合で再サンプリングすることによって解消することができる。これについては、以下の「顔モデルを構築するための例示システムと方法」の個所に詳しく説明されている。そこ10では、一定位置集合は「標準サンプル位置(standard sample positions)」と呼ばれている。再サンプリング関数はメッシュ変形関数(mesh deformation function)である。標準サンプル位置は、一度細分化された汎用メッシュの頂点に対応する、顔メッシュの頂点である。

【0036】第2の不一致では、ある顔から別の顔へ変位データを変換し、顔のサイズと形状の変化を補償する必要がある。図示の例では、これは、2つのデータセットについて対応する表情の小さなトレーニング集合を見付けたあと、一方から他方への最良線形変換を見付けることによって行われる。1つの例として、次の場合を考えてみる。すなわち、実験的環境では、感情を表す表情は、複数の表情のさまざまな強さを含む、49個の対応する表情に対して手作業でラベルが付けられていた。音声の動きについては、10,000フレームは、時間歪曲(time warping)手法を用いて自動的にアライメントされていた。

【0037】各表情は、m個の標準サンプル位置でのx,y,z変位のすべてを含んでいる3mベクトルgによって表されている。変換される顔のn個の表情ベクトル集合をg30 a1...anとし、ターゲット顔の対応するベクトル集合をgb1...bnとすると、線形予測子(linear predictors)集合ajは、下に示す3m線形最小二乗システムの解を求ることによって、gnの座標ごとに1つずつ計算される。

【0038】

【数3】

$$a_j \cdot g_{ai} = g_{bi}[j], i = 1..m$$

【0039】図示の例では、各ga_jのポイントは小さなサブセットだけが使用されている。具体的には、標準サンプルポイントとエッジを共有するポイントだけが考慮の対象になっている。使用されたメッシュでは、平均値(average valence)は約6であるので、代表的なga_jは18要素になっている。その結果のシステムは大体n x 18である。

【0040】結果の線形システムは条件付けが十分でない場合があり、その場合には、線形予測子ajは良好に概念化されない。特異値(singular value)の拡散は、擬似逆関数(pseudoinverse)を計算して、ajの解を求めると制御されるので、概念化が大幅に改善される。σ50 σ1未満の特異値はすべてゼロにされる。ただし、σ1は

マトリックスの最大特異値であり、 $\alpha = 0.2, \dots, 0.1$ である。

【0041】図5は、ここで説明している実施形態による表情変換方法のステップを記載しているフローチャートである。ステップ500で、表情のコードブックが用意される。この種のコードブックの例は上に示されている。ステップ502で、トレーニング表情集合が用意される。代表例として、このトレーニング集合は、コードブック表情が得られた人物とは異なる人物からの表情の集合になっている。トレーニング表情集合は、適切であれば、どの方法によっても取り込むことができる。1つの例として、表情は、図2に示すようなシステムを使用して取り込むことができる。トレーニング表情集合が用意されたあと、ステップ504で、トレーニング集合とコードブックを使用して変換関数が導き出される。この作業を行う1つの方法例は、上述した通りである。当然のことであるが、本発明の精神と範囲を逸脱しない限り、他の方法を使用することも可能である。例えば、ニューラルネットワーク(neural network)などの、種々の非線形変換を使用することも、基礎表情(basis expressions)の加重和を使用することも可能である。変換関数が導き出されると、これはコードブック内の表情のすべてに適用され、以後の顔アニメーション操作の基礎として利用できる合成表情集合が得られ、あるいは定義される。

【0042】例示のアプリケーション

図6はシステム600を示す。そこには、上述した表情変換プロセスをどのように採り入れるかが、1つの例で示されている。システム600は送信側コンピューティングシステム、つまり、トランスマッタ602と、受信側コンピューティングシステム、つまり、レシーバ604とを含み、これらは、インターネット(the Internet)などのネットワーク603を介して通信するように接続されている。トランスマッタ602は照明システム200(図2)を含み、これは、図2を参照して上述のように、ある人物の表情を取り込むように構成されている。さらに、トランスマッタ602は、図4を参照して上述したものと同じよう、コードブック400を含んでいる。なお、ここでは、コードブックは、上述したように合成表情集合に合成されているものとする。すなわち、照明システム200が取り込んだ表情をもつ人物から得られたトレーニング表情集合を使用して、コードブックが処理され、合成表情集合が得られている。

【0043】レシーバ604は再構築モジュール606を含み、これは、トランスマッタ602から受信したデータから顔画像を再構築するように構成されている。また、レシーバ604は、トランスマッタ602側のそれと同じコードブック400を含んでいる。ここでは、トランスマッタ602側にいる人物が、レシーバ604側にいる人物と通信しようとしている。トランスマッタ6

02側にいる人物が通信のために顔を動かすと、その顔表情と動きはトランスマッタ602によって取り込まれ、処理される。この処理には、顔表情を取り込み、合成コードブックをサーチして、コードブックの中で最も近似している合致表情を見付けることを含めることが可能である。合致する表情が合成コードブックの中で見つかると、その表情のインデックスがレシーバ604に送られ、アニメーション化された顔が再構築モジュール606を使用して再構築される。

【0044】例示の顔変換

図7は、ここで説明している実施形態による表情変換のいくつかの効果を示す図である。1行目の絵(A)は、第1人物(人物A)の合成顔を構成しており、そこには3つの異なる表情が示されている。これらの絵は、人物Aの取り込まれた顔の動きの結果である。第2人物(人物B)の顔の動きは、すでに取り込まれている。人物Bの取り込まれた顔の動きは3行目(C)に示されている。ここでは、3D動きデータは、いくつかのカラードットを人物の顔の上に置き、人物の顔が変形されたときドットの動きを測定することによって取り込まれているが、このことは、この分野の精通者ならば当然に理解されるものである。なお、顔の動きは、上述したシステムと方法によって取り込むことができる。人物Bの取り込まれた動きは、上述したように使用されて、人物Aの表情が変換されている。この操作の結果は、2行目に示されている。これら3セットの絵の表情は、すべて相互に一致している。同図に示すように、1行目(A)と2行目(B)の表情は、非常に異なる2人物から導き出されたとしても、非常に類似しているのに対し、第2人物(3行目(C))のオリジナル表情は、1行目(A)と2行目(B)の表情とは全体的に類似していない。

【0045】顔モデルを構築するための例示システムと方法

リアル感のある画像を作るために必要とされる顔のモデルは、2つの部分からなっている。モデルの第1部分は、顔のジオメトリ(つまり、顔面の形状)に関係している。他方、モデルの第2部分は、顔の反射率(つまり、顔のカラーと反射特性)に関係している。このセクションでは、このモデルの第1部分、つまり、顔のジオメトリについて説明する。

【0046】顔のジオメトリは、皮膚面に、目を表す別の面を加えたものから構成されている。この例では、皮膚面は、頭をレーザレンジでスキャンすることによって導き出され、変位マップをもつ細分面で表されている。目は別モデルであり、これは皮膚面と位置合わせされ、マージされて、高品質レンダリングに適した完全顔モデルが得られている。

【0047】メッシュ適合

顔モデルを構築する第1ステップでは、レーザレンジスキヤナで測定されたジオメトリに密接に近似する細分面

(12)

21

が作成される。図示の例では、細分面は、Loopの細分ルールを使用して粗三角メッシュから定義されている。Loopの細分ルールは、Charles Loop著「三角形に基づく平滑細分面(Smooth Subdivision SurfacesBased on Triangles)」(PhD thesis, University of Utah, August 1987)に詳しく説明されている。さらに、細分面は、Hoppe他著「区分平滑面再構築(Piecewise smooth surface reconstruction)」(Computer Graphics (SIGGRAPH '94 Proceedings) pp. 295-302, July 1994)に記載されているものと類似のシャープエッジを追加して定義されている。なお、不規則なクリーズマスク (crease mask) は使用されていない。さらに、ダーツ(dart)とクリーズ頂点(crease vertex)の間のエッジを細分化するとき、クリーズ頂点に隣接する新エッジだけがシャープエッジとしてマークされている。

【0048】すべての顔モデルに対する細分面は、シングルベースメッシュを使用して定義され、頂点位置だけが、各々の異なる顔の形状に適応するように変化している。図示の例では、227個の頂点と416個の三角形をもつベースメッシュは、顔の全体形状をもち、目と唇付近が詳細化されるように定義されており、そこでは、最も複雑なジオメトリと動きが現れている。口の開きはメッシュの境界になっており、上唇と下唇上の対応する頂点の位置を1つに結ぶことによって、適合プロセス期間は、閉じたままになっている。このベースメッシュには、シャープな細分ルール用にマークされた少數のエッジがあり、これらは、口の開きの両側にコーナを作成し、鼻の両側を折り重ねるときの場所を得るのに役立っている。修正された細分ルールでは、少なくとも3つのシャープエッジのチェインに対してだけクリーズを取り入れられているので、このモデルは表面にクリーズがない。孤立頂点(isolated vertex)だけは、明確に定義された限界法線(limit normal)をもつことができないからである。

【0049】図8は、この例に従って使用された粗定義メッシュ(中央の図)の例を示す。図8は、この粗メッシュをどのように使用して、同じ細分制御(粗)メッシュを各々の顔の変位細分面にマッピングすると、ある顔と別の顔との間に自然な対応関係が得られるかを示したものである。この側面は、以下に詳しく説明されている。

【0050】細分面を各顔に適合させるために使用されるプロセスは、Hoppe他著「区分平滑面再構築(Piecewise smooth surface reconstruction)」(Computer Graphics (SIGGRAPH '94 Proceedings) pp. 295-302, July 1994)に記載されているアルゴリズムに基づいている。Hoppeの面適合法は、基本的には、3つのフェーズからなるものと言うことができる。すなわち、トポロジ型推定(フェーズ1)、メッシュ最適化(フェーズ2)、および区分平滑面最適化(フェーズ3)である。

特開2002-123837

22

【0051】フェーズ1では、未知の面上または付近に未組織のポイント集合が与えられている、相対的に多数の三角形からなる三角メッシュが構築される。このフェーズでは、面のトポロジ型が判断され、ジオメトリの初期推定が得られる。フェーズ2は、フェーズ1の出力から始まり、三角形の数を減らし、データへの適合を改善する。このアプローチでは、表現の簡潔化と適合の良好化という、相互に競合する目標間のトレードオフを明示的にモデル化するエネルギー関数を最適化することを課題としている。この最適化手順における自由変数としては、メッシュ内の頂点の数、その接続性(connectivity)、およびその位置がある。フェーズ3は、フェーズ2で得られた最適化メッシュ(区分線形面)から出発し、この場合も、簡潔性とデータへの適合性をトレードオフするエネルギー関数を最適化することによって、正確で、簡潔な区分平滑細分面を適合させる。フェーズ3で最適化が行われると、制御メッシュ内の頂点の数、その接続性、その位置、およびシャープな特徴の数とロケーションが変化する。面におけるシャープな特徴の自動検出と回復は、このフェーズの不可欠な部分となっている。

【0052】この実施形態では、処理の仕方が、Hoppe他に記載されているアプローチとは、2つの点で異なっている。第1は、制御メッシュの接続性を変更する必要がないので、連続的最適化は、頂点位置に対してだけ行われることである。さらに、特徴の制約が、平滑化項と共に追加されていることである。

【0053】図示の例では、適合プロセスは次の汎関数を最小化する。

【0054】

【数4】

$$E(\mathbf{v}) = E_d(\mathbf{v}, \mathbf{p}) + \lambda E_r(\mathbf{v}) + \mu E_t(\mathbf{v})$$

【0055】上記において、 \mathbf{v} はすべての頂点位置のベクトル、 \mathbf{p} はレンジスキナからの全データポイントのベクトルである。3つの項の下付き文字は、距離、形状、および制約を表している。距離関数 E_d は、レンジスキナポイントから細分面までの和二乗距離(sum-squared distance)を測定する。

【0056】

【数5】

$$E_d(\mathbf{v}, \mathbf{p}) = \sum_{i=1}^{n_p} a_i \| \mathbf{p}_i - \Pi(\mathbf{v}, \mathbf{p}_i) \|^2$$

【0057】上記において、 \mathbf{p}_i は i 番目のレンジポイント、 $\Pi(\mathbf{v}, \mathbf{p}_i)$ は、頂点位置 \mathbf{v} で定義された細分面上のレンジポイントの投影である。重み a_i は、 \mathbf{p}_i におけるスキナの視点方向が $\Pi(\mathbf{v}, \mathbf{p}_i)$ における面法線と一致していないポイントを無視させるブール項である。さらに、面からある距離以上に離れているポイントは拒否される。

【0058】

【数6】

$$a_i = \begin{cases} 1 & \text{if } \langle s(p_i), n(\Pi(v, p_i)) \rangle > 0 \text{ and } \|p_i - \Pi(v, p_i)\| < d_0 \\ 0 & \text{otherwise} \end{cases}$$

【0059】上記において、 $s(p)$ は、ポイント p におけるスキャナの視点に向かう方向、 $n(x)$ はポイント x における外側に面する面法線である。

【0060】平滑関数 E_s は、制御メッシュをローカルに平坦化させるものである。この関数は、各頂点から隣接頂点の平均までの距離を測定する。

【0061】

【数7】

$$E_s(v) = \sum_{j=1}^{n_v} \left\| v_j - \frac{1}{\deg(v)} \sum_{i=1}^{\deg(v)} v_i \right\|^2$$

【0062】頂点 v_{ki} は、 v_j の隣接頂点である。

【0063】制約関数 E_c は、単純には、制約頂点集合から対応するターゲット位置集合までの和二乗距離である。

【0064】

【数8】

$$E_c(v) = \sum_{i=1}^m \|A_c v - d_i\|^2$$

【0065】上記において、 A_j は、 j 番目の頂点の限界位置を、制御メッシュからとらえて定義し、頂点 c_i の限界位置が 3D ポイント d_i に付加されるようにする線形関数である。制約は、最適化変数の再パラメータ化によって厳格に適用することも可能であるが、この制約をソフトにすると、繰り返しが望ましいローカル最小限まで円滑に行われることが判明している。制約は、汎用メッシュの顔特徴を、適合される特定の顔上の、対応する特徴に合致させるためにユーザによって選択される。この例では、約 25 乃至 30 の制約が使用され、目、鼻、および口に集中している。図 9 は、900 での細分制御メッシュ (A) の制約と、顔モデル (B) 上の対応するポイントを示す。

【0066】 $E(v)$ を最小限にすることは、 Π と a_i が v の線形関数でないので、非線形最小二乗問題である。しかし、 a_i を一定に保ち、 $\Pi(v, p_i)$ を一定の線形的制御頂点の組み合わせによって近似化すると、これを線形問題にすることができる。従って、適合プロセスは、線形最小二乗問題のシーケンスとして進められ、 a_1 および面上の p_i の投影は、各繰り返しの前に再計算されることになる。細分限界面は、これらの計算を行うために、特定の細分レベルにあるメッシュによって近似化される。顔を適合することは、少數の繰り返しで行われ（30 未満）、制約は、繰り返しの進行と共に、単純なスケジュールに従って更新されて行く。その更新は高 λ と低 μ か

ら始まって、最適化を非常にスムーズな顔の近似化に導いて行き、低 λ と高 μ まで進んで行き、最終的解により、データと制約は密接に適合することになる。実際の計算時間は、 $\Pi(v, p_i)$ の計算時間によって決まる。

【0067】レンダリングのためのメッシュを作るため 10 に、面は望ましいレベルまで細分化され、顔の形状にスムーズに近似しているメッシュが得られる。変位は、各々の頂点ごとに、その頂点での面に対するライン法線を、オリジナルスキャンで定義された三角面と交差させることによって計算される。これについては、Lee 他著「変位細分面 (Displaced Subdivision Surfaces)」(SIGGRAPH '00 Proceedings, July 2000) に説明されている。その結果の面は、オリジナルスキャンの目立った特徴をすべて、若干三角形が少なくなっているメッシュに再現しているが、これは、ベースメッシュは顔の重要な領域 20 の三角形が多くなっているためである。この細分ベースの表現によると、面をパラメータ化したものと、そのパラメータ化で定義された組み込み多重解像基底関数集合 (built-in set of multiresolution basis functions) とが得られることになり、さらに、特徴制約が適合時 使用されるため、この方法を使用して適合されるすべての顔にわたって自然の対応関係が得られることになる。この構造は、顔のアニメーションで利用すると、多くの面で便利である。

【0068】図 10 は、ここで説明している実施形態に 30 従って顔モデルを構築するための方法のステップを記載しているフローチャートである。この方法は、適当なハードウェア、ソフトウェア、ファームウェアまたはこれらの組み合わせで実現することが可能である。この例では、方法はソフトウェアで実現されている。

【0069】ステップ 1000 で、1 つまたは 2 つ以上の顔の 3D データが測定され、対応する顔モデルが得られる。上記の例では、3D データは、顔をレーザレンジでスキャンすることによって得られていた。当然に理解されるように、3D データは他の方法で得ることも可能である。ステップ 1002 で、1 つまたは 2 つ以上の顔モデルに適合するために使用される汎用顔モデルが定義される。当然に理解されるように、汎用顔モデルは、多数の異なる顔に適合するように利用できるという利点がある。従って、このことは、そのようなことが行われていなかった従来の方法に比べて、改良点となっている。上述した例では、汎用顔モデルは、粗三角メッシュ (coarse triangle mesh) の形体をしたメッシュ構造になっている。この三角メッシュは、顔のジオメトリに密接に近似している細分面を定義している。図示の例では、顔モデルのすべてについての細分面は、シングルベースメッシュ

(14)

25

シユを使用して定義されている。ステップ1004で、汎用顔モデル上の個々のポイントまたは制約が選択される。これらの個々のポイントまたは制約は、顔モデル上にマークされた、対応するポイントに直接にマッピングされる。これら個々のポイントのマッピングは、起こり得る多数の、異なる顔モデルの各々について、同じように行われる。ステップ1006で、汎用顔モデルは1つまたは2つ以上の顔モデルに適合される。このステップは、各々の異なる顔の形状に適応するように、頂点の位置だけを操作することによって行われる。適合プロセス期間には、頂点位置だけにわたって連続的最適化が行われるので、メッシュの連続性が変更されることはない。さらに、この適合プロセスでは、個々のポイントまたは制約が直接に顔モデルにマッピングされる。さらに、平滑項 (smoothing term) が追加され、これは、制御メッシュがローカルに平坦化されるように最小化されている。

【0070】目の追加

上述してきた変位細分面は、顔の皮膚面の形状を非常によく表している。しかし、リアル感のある顔を得るために望ましいとされる特徴が、他にもいくつかある。その中で最も重要な特徴が目である。レザレンジスキナは目にに関する適切な情報を取り込まないので、メッシュは、別々にモデル化された目を追加することによって、レンダリングのために補強されている。顔モデルの他の部分と異なり、目とその動きは、特定の人物から測定されないので、リアルな目の外観を再現するとは限らないようになっている。しかし、目の存在と動きは、顔モデルの全体的外観にとって非常に重要である。

【0071】適切な目モデルであれば、その目モデルを使用して目をモデル化することができる。図示の例では、2部分からなるモデルを構築するために、市販のモデリングパッケージが使用されている。第1部分は眼球のモデルであり、第2部分は目回りの皮膚面のモデルであり、その中には、まぶた、眼窓、および周囲顔面部分が含まれている（この第2部分は「眼窓面 (orbit surface)」と呼ぶことにする）。目を全体顔モデルの一部にするためには、眼窓面を、モデル化しようとする個人の顔に適合させる必要があり、2つの面を縫い合わせる必要がある。これは2ステップで行われる。最初に、2つのメッシュは、眼窓面上に定義された重み付け関数に従ってゆがみが付けられ、顔と眼窓が重なり合った個所で同時に現れるようにする。次に、2つの面はペアの同心梢円面でカットされ、1つのメッシュになるように縫い合わされる。

【0072】なお、上述してきた実施形態の利点のある特徴の1つは、これらの実施形態によると、ある人物の表情を別の人物の表情に変換するために使用できる構造またはフレームワークが得られることである。汎用顔モデルを各個人の顔に適合することは、汎用モデルのどの

特開2002-123837

26

部分も、各人物の顔上の同じ特徴に常にマッピングされるように制約されている。例えば、汎用モデル中の口の左隅は、どの人物の顔でも、その口の左隅に常にマッピングされる。適合顔モデルの集合を利用すると、どの顔においても、ある特定の顔にある特定のポイントに対応するポイントを判断することができる。例えば、人物Aの顔上の口の左隅の動きが測定されたとする。顔Aへの汎用モデルの適合を使用すると、汎用モデルのどのポイントが測定ポイントに対応しているかを判断することができ、次に、顔Bへの汎用モデルの適合を使用すると、Bの顔のどのポイントが汎用モデル上の、計算されたポイントに対応しているか、従って、顔Aの測定ポイントに対応しているかも判断することができる。測定した顔上のポイント集合から動きを再現するために新しい顔のどの部分を移動させる必要があるかを知っている必要があるため、この情報は、動きをある顔から別の顔に変換する上で不可欠である。

【0073】顔の動き

顔の動きは、画面上のサンプルポイント集合の、時間と共に変化する3D位置によって指定される。顔が動き取り込みデータによって制御されるとき、これらのポイントは、動き取り込みシステムによってトラッキングされる、顔上のマーカになっている。これらのポイントの動きは、顔面の領域にスムーズに影響する制御ポイントの集合によって顔面を制御するために使用される。顔の動きデータの取り込みは、この分野の精通者ならば理解されるように、いずれかの適当な方法で行うことができる。1つの具体例では、顔の動きは、Guenter他著「顔の作成(Making Faces)」(Proceedings of SIGGRAPH 1998, pp. 55-67, 1998) に記載されている手法を使用して取り込まれていた。

【0074】メッシュ変形

顔は、制御ポイント集合 q_j の変位の線形的組み合わせに従って、三角メッシュの各頂点 w_i をその静止位置から変位することによってアニメーション化されている。これらの制御ポイントは、動きを記述しているサンプル点 p_j と1対1で対応している。各制御ポイントが頂点に及ぼす影響は、対応するサンプルポイントからの距離と共に低下していく、複数の制御ポイントが頂点に影響する場合には、それらの重みは総和が1になるように正規化される。

【0075】

【数9】

$$\Delta w_i = \frac{1}{\beta_i} \sum_j \alpha_j \Delta q_j ; \alpha_j = h(\|w_i - p_j\| / r)$$

【0076】上記において、頂点 i が複数の制御ポイントによって影響される場合、 $\beta_i = \sum_k \alpha_{ik}$ であり、そうでなければ1である。これらの重みは、サンプルポイントと顔メッシュの静止位置を使用して一回計算される

ので、フレームごとのメッシュの動きはスパースマトリックス乗算 (sparse matrix multiplication) になっている。重み付け関数については、次式が使用されている。

【0077】

【数10】

$$h(x) = \frac{1}{2} + \frac{1}{2}\cos(\pi x).$$

【0078】顔をアニメーション化する細部を処理するには、これらの重み付けルールには2つタイプの例外が設けられている。目と口の付近の頂点と制御ポイントは、「上」および「下」のタグが付けられ、例えば、口の上有る制御ポイントは口の下にある頂点の動きには影響しない。また、目の周囲の領域におけるスカラテクスチャマップ (scalar texture map) は、目がスムーズなテープ状になって、まぶたの個所でゼロになるように動きに重み付けするために使用されている。サンプルポイント集合に従って顔メッシュを動かすには、制御ポイント位置を計算して、顔面がそれなりに変形されるようにならなければならない。上述したのと同じ重み付け関数を使用すると、サンプルポイントが制御ポイントに応じてどのような動きをするかを計算することができる。その結果は線形変換 $p = Aq$ である。従って、時刻 t に、サンプル位置 p_t を求めるには、制御ポイント $q_t = A^{-1}p_t$ を使用することができる。しかし、マトリックス A は条件付けが十分でないことがあるので、制御ポイントの動きが非常に大きいために起こる、望ましくない面形状を回避するために、SVD (Singular Value Decomposition—特異値分解) を使用して A^{-1} が計算され、 A^{-1} の特異値は限界 M でクランプされる。図示の例では、 $M = 1.5$ が使用されている。SVDを論じている参考文献としては、Golub および Van Loan著「マトリックス計算 (Matrix Computations)」第3版、Johns Hopkins press, 1996がある。

【0079】目と頭の動き

顔を実物に近い外観にするためには、プロシージャで生成された動きが目に加えられ、別々に取り込まれた剛体運動が頭全体に加えられる。眼球は、ランダムな固定方向のシーケンスに従って回転され、一方の眼球から次の眼球に向かってスムーズに移動して行く。まぶたは、眼球の中心を通る軸周りにまぶたを定義している頂点を回転し、まぶたメッシュ上に定義された重みを使用して変形がスムーズに行われるようによることでアニメーション化される。

【0080】頭の剛体運動は人物の頭の物理的動きから取り込まれ、これは、その人物が特殊なマシン認知可能ターゲットでマークされた帽子を着用している間にその動きをフィルムに写すことによって行われる（帽子は、Marschner他著「人間の皮膚を含む画像ベースBRDF測定 (Image-based BRDF measurements including human skin)」、Rendering Techniques'99 (Proceedings of the Eurographics Workshop on Rendering), pp. 131-144, June 1999に記載されているものがある。これらの測定は、いくつかの被写体の額の平均BRDFを記述しており、Lafontaine他著「反射率関数の非線形近似化 (Non-linear approximation of reflectance functions)」、Computer Graphics (SIGGRAPH '97 Proceedings), pp. 117-126, August 1997に記載されているBRDFモデルの適合パラメータを含んでいる。従って、これらの測定は、リアル感のある顔をレンダリングするときの、すぐれた出発点となっている。しかし、これらの測定は、実際の顔で観察される空間的変化の一部を含むように補強する必要がある。これは、その皮膚がレンダリングした皮膚と類似している被写体の測定BRDFへの適合から始め、それ

Eurographics Workshop on Rendering), pp. 131-144, June 1999で使用されているものに近似するようにパターン化されている）。これらのターゲットをビデオシークエンスでトラッキングすることにより、頭の剛体運動が計算され、その後、これは頭モデルをレンダリングするために適用される。このセットアップによると、ビデオカメラだけで済むので、望みのアクションを実演しながら頭の動きを作っていくことができる。

【0081】反射率をモデリングするための例示システムと方法

リアル感のある顔画像をレンダリングするためには、ジオメトリが正確であるだけでなく、皮膚からの光の反射の計算が正確であることが要求される。図示の例では、顔をレンダリングするために、物理ベースの Monte Carlo 光線トレーサが使用されている。手法の例としては、Cook他著「分布光線トレーシング (Distribution Ray Tracing)」、Computer Graphics (SIGGRAPH '84 Proceedings), pp. 165-174, July 1984 および Shirley他著「直接照明計算のための Monte Carlo 手法 (Monte Carlo techniques for direct lighting calculations)」、Transactions on Graphics, 15(1):1-36, 1996に記載されているものがある。この方法によると、任意の BRDF (bi-directional reflectance distribution functions: 双方向反射率分布関数) を使用して、単純なシェーディングモデルでは近似化が困難であった、皮膚の外観を正確にシミュレートすることができる。さらに、拡張光源が使用されているが、これは、ポートレート写真のようにレンダリングするとき、見苦しくない画像を得るために必要になるものである。物理的光移動は、計算効率の必要上、次の2点でずれるようにしている。すなわち、拡散相互反射 (diffuse interreflection) は無視され、目は屈折なしで瞳を通るように照明されている。

【0082】図示の例では、皮膚の反射率モデルは、実際の人間の顔の測定に基づいている。手法の例としては、Marschner他著「人間の皮膚を含む画像ベースBRDF測定 (Image-based BRDF measurements including human skin)」、Rendering Techniques'99 (Proceedings of the Eurographics Workshop on Rendering), pp. 131-144, June 1999に記載されているものがある。これらの測定は、いくつかの被写体の額の平均BRDFを記述しており、Lafontaine他著「反射率関数の非線形近似化 (Non-linear approximation of reflectance functions)」、Computer Graphics (SIGGRAPH '97 Proceedings), pp. 117-126, August 1997に記載されているBRDFモデルの適合パラメータを含んでいる。従って、これらの測定は、リアル感のある顔をレンダリングするときの、すぐれた出発点となっている。しかし、これらの測定は、実際の顔で観察される空間的変化の一部を含むように補強する必要がある。これは、その皮膚がレンダリングした皮膚と類似している被写体の測定BRDFへの適合から始め、それ

を拡散反射成分と鏡面反射成分に分離することによって行われる。そのあと、各々をモジュレートするためにテクスチャマップが取り入れられている。

【0083】拡散反射成分のテクスチャマップ、つまり、「アルベドマップ(albedo map)」は、以下で説明するように、被写体の実際の顔から取られた測定に従つて、拡散反射率をモジュレートする。鏡面反射成分は、スカラテクスチャマップ(scalartexture map)によってモジュレートされ、皮膚反射率でレンダリングしてはならない個所(まゆやまつ毛など)から鏡面反射が除去され、顔の下部の鏡面反射が低減され、顔の皮膚の特徴に近似するようにされる。その結果は、空間的に変化するBRDFであり、これは、Lafortune他著「反射率関数の非線形近似化(Non-linear approximation of reflectance functions)」、Computer Graphics (SIGGRAPH '97 Proceedings), pp. 117-126, August 1997の概念化コサインローブ(generalized cosine lobes)の和によって各ポイントに記述されている。

【0084】アルベドマップの構築

図示し、上述してきた実施形態では、拡散反射に起因する、空間的に変化する反射率を記述していることが要求されるアルベドマップは、制御された照明の下で撮影された顔のデジタル写真シーケンスを使用して測定されていた。

【0085】図11は、デジタル写真または画像を取り込むために使用された例示システムを示している。図示のシステムでは、デジタルカメラ1100が装備され、複数の光源を含んでいる。そのいくつかの例は1102、1104で示されている。垂直ポラライザ1106、1108、および1110の形体をした偏光フィルタが装備され、光源とカメラレンズを被覆しているため、鏡面反射が抑止され、拡散反射成分だけが画像に残るようにしている。この例では、被写体は、マシン認識可能ターゲットがプリントされた帽子1112を着用し、頭のポーズをトラッキングするようにしている。カメラ1110は、被写体が回転している間静止したままになっている。唯一の照明は、カメラ付近の測定個所に置かれた光源1102、1104から得ている。漏れた光からの間接的反射を減らすため、黒のバックドロップ(背景幕)が使用されている。

【0086】カメラと光源のロケーションは分かっているので、面法線、放射照度(irradiance)、視点方向、およびテクスチャ空間内の対応する座標は、各画像のピクセルごとに標準的光線トレーシング手法を使用して計算することができる。理想的なランベルト(Lambertian)反射が観察されるとの想定の下では、ランベルト反射率は、テクスチャ空間内の特定のポイントについてこの情報から計算することができる。この計算は、1つの写真に含まれるピクセルごとに繰り返され、その総和が得られると、画像がテクスチャ空間内に投影され、光源

に起因する計算された放射照度で除すると、面を横切る拡散反射率のマップが得られる。図12に示すように、2つの写真(A)、(B)がテクスチャ空間内に投影され、照明が補正されている。実際には、この投影は、外側ループがテクスチャマップ内のすべてのピクセルを通るように繰り返される逆マッピング(reverse mapping)によって行われ、特定のテクスチャピクセルに投影される画像エリア全体の平均をとるために、確率的スーパサンプリング(stochastic supersampling)が使用されている。

【0087】単一の写真からのアルベドマップは、面の一部だけをカバーしており、その結果は、グレージング(grazing)角度が小さいとき最良になっている。従つて、顔全体のシングルアルベドマップを作成するために、すべての個別マップの加重平均が計算されている。重み付け関数は、そのビジュアル例が図13(A)、(B)に示されているが、面に対してほぼ法線をなしている方向から表示および/または照射されるピクセルには高い重みが与えられるように選択され、表示か照射のどちらかが極度にグレージングする前にゼロに降下するようになっている。図示の例では、次の関数が使用されている。 $(\cos \theta_i \cos \theta_e - c)^p$ 、ただし、 $c = 0.2$, $p = 4$ 。

【0088】特定のテクスチャピクセルのアルベドを計算する前に、ピクセルが可視であり、適当に照射されていることが確かめられる。複数の光線は、そのピクセル上のポイントから光源上のポイントまでとカメラポイントまでトレースされ、そのピクセルには、可視性と照明がゼロであるか、部分的であるか、完全であるかのマークが付けられる。光源のサイズを見積もるとき、重大なエラーを避けるのが賢明である。完全に可視であり、少なくとも1つの光源によって完全に照射され、どの光源によつても部分的に照射されないピクセルのアルベドだけが計算される。このようにすると、部分的に隠されたピクセルおよび全影または半影領域にあるピクセルは使用されることがない。

【0089】これらの測定に意味をもたせるためには、若干の較正が必要である。カメラの転写曲線は、Debevec他著「写真からの高ダイナミックレンジ放射輝度マップの回復(Recovering high dynamic range radiance maps from photographs)」、Computer Graphics (SIGGRAPH '97 Proceedings), pp. 369-378, August 1997に記載されている方法を使用して較正されている。光/カメラシステムのフラットフィールド応答は、大きなホワイトカードの写真を使用して較正されている。レンズのオーバカルレンジスとひずみは、Zhang著「カメラ較正のためのフレキシブルな新手法(A flexible new technique for camera calibration)」、Technical Report MSR-TR-98-71, Microsoft Research, 1998に記載されている手法を用いて較正されている。全体スケールファクタは、

(17)

31

既知反射率の参照サンプルを使用してセットされている。光源強度の画像間の変化が考慮の対象となっているときは、すべての画像に参照サンプルを組み込むことによって制御が行われている。

【0090】上記プロセスから得られたテクスチャマップは、顔全体にわたる詳細なカラー変化を自動的に取り込む上で良好な働きをしている。さらに、較正帽子を定位位置に保持するために使用されるストラップは可視になっている。これらの問題は、画像編集ツールを使用し、プランクエリアに近傍テクスチャまたは均一カラーを埋め込むことによって解消されている。

【0091】図14(A), (B)と図15(A), (B)は、編集前のアルベドマップと編集後のアルベドマップを、比較目的で示したものである。アルベドマップから妥当な結果が得られない個所がどこであるかは、面が十分に観察されない個所(例えば、あごの下)であるか、あるいは面が余りに複雑な形状をしているため画像と共に正しくスキャンし、記録できない個所(例えば、耳)であることは明らかである。これらの個所は、どちらも、リアル感のある外観を得るためにアルベドマップからのテクスチャを必要としていない。第1の個所は視覚的に目立たないからであり、第2の個所は、ジオメトリがビジュアル的に詳細になっているからである。そのため、これらの個所を編集しても、最終的レンダリングの外観はほとんど影響されない。

【0092】図16～図18は、付随するビデオからの静止フレームを使用した顔モデルのいくつかの、異なる様相を示す図である。1行目(図16の(A), (B), (C))には、いくつかの角度からの顔が示されている。そこには、アルベドマップと測定されたBRDFが、任意の角度から見た顔全体にわたる皮膚の外観とそのカラー変化をどのように取り込むかが示されている。

2行目(図17の(A), (B), (C))は、グレージング角度での強い鏡面反射を含む、リムとサイドの照明の効果を示している。なお、光源は、この行の3画像のいずれも、同じ強度であり、顔から同じ距離に置かれている。反射率の方向変化は、親しみのある照明効果がレンダリングに現れることに貢献している。3行目(図18の(A), (B), (C))では、表情の変形が顔に加えられても、顔は通常の表情の動きでは自然のように見えることを示している。

【0093】図19は、ここで説明している実施形態に従ってアルベドマップを作成するための方法のステップを記載しているフローチャートである。この方法は、適切なハードウェア、ソフトウェア、ファームウェアまたはこれらの組み合わせで実現することが可能である。ここで説明している実施形態では、この方法は、図11を参照して説明してきたシステムと同じシステムで使用されるものとして、ソフトウェアで実現されている。

【0094】ステップ1700で、被写体を照射するた

特開2002-123837

32

めに使用できる1つまたは2つ以上の偏光光源が準備される。光源の例は上述した通りである。ここで説明している実施形態では、光源は、被写体の顔反射率の鏡面反射成分が抑止または除去されるように選択されている。ステップ1702で、被写体の顔は光源で照射される。ステップ1704で、被写体を回転させ、その間に一連のデジタル写真または画像が取り込まれる。ステップ1706で、テクスチャマップ内のピクセルごとに、面法線、放射照度、視点方向およびテクスチャ空間内の座標が計算される。この計算は公知のアルゴリズムを使用して行われる。ステップ1708で、画像のテクスチャ空間内の特定ピクセルに対するランベルト反射率が計算される。これにより、ピクセルのアルベドが得られる。ステップ1712で、追加のピクセルがアルベドマップに残っているかどうかが判断される。残っていれば、ステップ1712で次のピクセルが得られ、ステップ1708に戻る。追加のピクセルがアルベドマップに残っていないければ、ステップ1714で、追加のデジタル画像が残っているかどうかが確かめられる。追加のデジタル画像が残っていれば、ステップ1716で次のデジタル画像が得られ、ステップ1706に戻る。追加のデジタル画像が残っていないければ、ステップ1718で、各画像の個別アルベドマップの加重平均が計算され、顔全体のシングルアルベドマップが作成される。この加重平均処理がどのように行われるかの具体例は、Marschner著

「コンピュータグラフィックスの逆レンダリング(Inverse Rendering for Computer Graphics)」、PhD thesis, Cornell University, August 1998に記載されている。

【0095】図20は、シングルピクセルのアルベドを計算する方法のステップを記載しているフローチャートである。この方法は、適切なハードウェア、ソフトウェア、ファームウェアまたはこれらの組み合わせで実現することが可能である。ここで説明している実施形態では、この方法はソフトウェアで実現されている。ステップ1800で、あるピクセルが与えられているとき、そのピクセルが完全に可視であるかどうかが判断される。ピクセルが完全に可視でなければ、ピクセルのアルベドは計算されない(ステップ1804)。ピクセルが完全に可視であれば、ステップ1802で、ピクセルが少なくとも1つの光源で完全に照射されているかどうかが判断される。ピクセルが少なくとも1つの光源で完全に照射されていなければ、ピクセルのアルベドは計算されない(ステップ1804)。ピクセルが少なくとも1つの光源で完全に照射されていれば、ステップ1806で、ピクセルがいずれかの光源で部分的に照射されているかどうかが判断される。そうであれば、そのピクセルについては、アルベドは計算されない。ピクセルがいずれかの光源で部分的に照射されていなければ、ステップ1808で、そのピクセルのアルベドと重みが計算される。この重みは、あとで個別マップをまとめて、その平均を

とるとき使用される。従って、上述したように、アルベドが計算されるのは、完全に可視であり、少なくとも1つの光源で完全に照射され、いずれかの光源で部分的に照射されていないピクセルについてだけである。このようにすると、部分的に隠されているピクセルと全影または半影領域にあるピクセルは使用されることがない。

【0096】以上、説明した顔の特徴をアニメーション化する方法および表情変換のための方法を実現するためのアプリケーションプログラムは当業者であれば、本明細書の説明に基づいて、当業者であれば作成することが可能であろう。

【0097】アプリケーションプログラムはCPUが実行可能な読み取り実行可能な命令で作成すればよい。また、上記アプリケーションプログラムを記憶しておく記録媒体は、メモリ、ハードディスク、周知の携帯用記録媒体に記憶すればよく、これらの記録媒体は1つまたは2つ以上でもよい。たとえば、アプリケーションで使用するデータ、たとえば、コードブックやトレーニング表情を定義したデータ等はCD-ROMに記憶しておき、これらのデータを使用するアプリケーションプログラムはハードディスクに記憶しておくという異なる記録媒体にアプリケーションプログラムとデータとを記憶させてもよい。また、同一の記録媒体に記録させててもよいこと勿論である。

【0098】さらにアプリケーションプログラムを実行するコンピュータは1台でもよいしアプリケーションプログラムで規定した処理を複数台のコンピュータで分散処理してもよい。

【0099】

【発明の効果】結論

上述してきた実施形態によれば、顔をリアル感のある顔にモデリングし、レンダリングするという高い基準を満たしてから、人間の顔と同じように親しみのある画像が信頼できる形で表示できるようにするという難題を解決するシステムと方法が提供されている。このアプローチの考え方には、顔モデルが実際に現実の顔に似るように、可能な限り測定を利用するにある。顔のジオメトリは、異なる顔に共通して統一した接続性と対応性をもつ、変位でマッピングされた細分面で表されている。反射率は、従来による人間の皮膚のBRDF測定から得られ、その測定が新規な測定と併用されて、複数の視点（ビュー）が拡散反射率を表す、単一の照明補正テクスチャマップになるように結合されている。動きは、上述した動き取り込み手法によって得られ、その動きが、改良された変形方法を使用して顔モデルに適用されることによって、シェーディング面に適した動きが得られるようにしている。レンダリングのリアル感は、現実の顔のジオメトリ、動きおよび反射率を物理ベースのレンダラー（physically-based renderer）で使用することによって大幅に向上されている。

【0100】以上、構造上の特徴および/または方法ステップからとらえて本発明を説明してきたが、以上の説明から理解されるように、請求項に明確化されている本発明は、上述した具体的特徴または方法ステップに必ずしも限定されるものではない。むしろ、具体的特徴および方法ステップは、請求項に記載の発明を実現する、好ましい実施形態として開示されたものである。

【図面の簡単な説明】

【図1】本発明の実施形態を実現する上で使用するのに適している汎用コンピュータを示すブロック図である。

【図2】被写体の顔の構造情報と反射率情報を同時に取り込むために利用できるシステムを示す構成図である。

【図3】本発明の実施形態に従って構造情報と反射率情報を取り込むための例示方法を記載しているフローチャートである。

【図4】本発明の実施形態による例示コードブックと変換関数を示す説明図である。

【図5】本発明の実施形態による表情変換プロセスを示すフローチャートである。

【図6】本発明の実施形態のある種の原理が採用可能である例示システムを示すブロック図である。

【図7】(A)～(C)は本発明の実施形態による例示表情変換を示す例示カラープレートのコレクションを示す図である。

【図8】同一细分制御メッシュを、顔が異なるごとに変位された细分面にマッピングするプロセスを示す図である。

【図9】(A)および(B)は面適合過程で特徴の対応づけを行うために利用される例示制約を示す図である。

【図10】本発明の実施形態による面適合方法のステップを記載しているフローチャートである。

【図11】本発明の実施形態に従って顔のアルベドマップを作るために採用できる例示システムを示す構成図である。

【図12】(A)および(B)はテクスチャ空間内に投影され、照明が補正された2つの写真の例示アルベドマップを示す図である。

【図13】(A)および(B)は図12に示す写真に対応する例示重み付け関数を示す図である。

【図14】(A)および(B)は2つの異なるデータセットに対する2つの完全アルベドマップを示す図である。

【図15】(A)および(B)は編集したあとの図14のアルベドマップを示す図である。

【図16】(A)～(C)は異なる照明条件の下で異なる向きにレンダリングされた顔モデルのコレクションを示す図である。

【図17】(A)～(C)は異なる照明条件の下で異なる向きにレンダリングされた顔モデルのコレクションを

示す図である。

【図18】(A)～(C)は異なる照明条件の下で異なる向きにレンダリングされた顔モデルのコレクションを示す図である。

【図19】本発明の実施形態に従ってアルベドマップを作成するための方法のステップを記載しているフローチャートである。

【図20】本発明の実施形態に従ってシングルピクセルのアルベドを計算するための方法のステップを記載しているフローチャートである。

【符号の説明】

200 照明システム

202 カメラ

202a カメラのレンズ

204、206 光源

400 コードブック

402 トレーニング表情集合

406 変換プロセッサ

408 変換関数

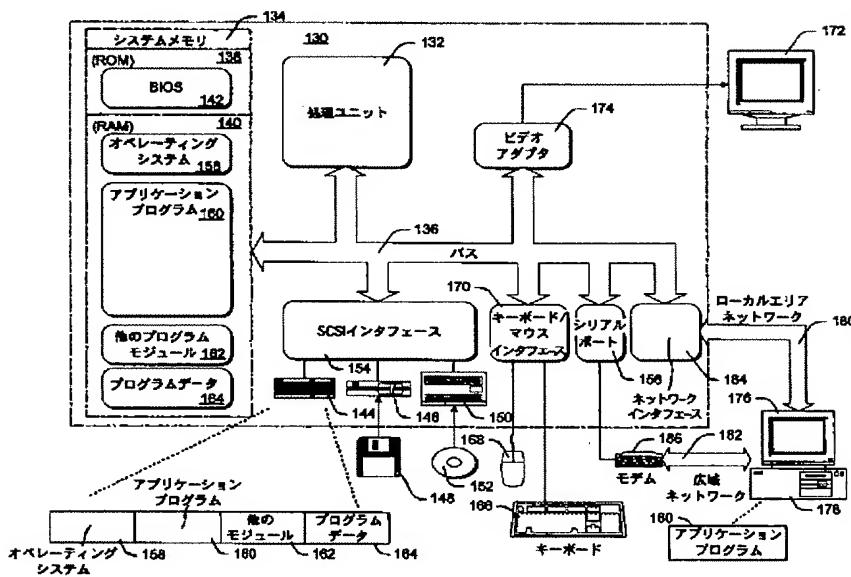
410 合成表情集合

602 トランスマッタ(送信側)

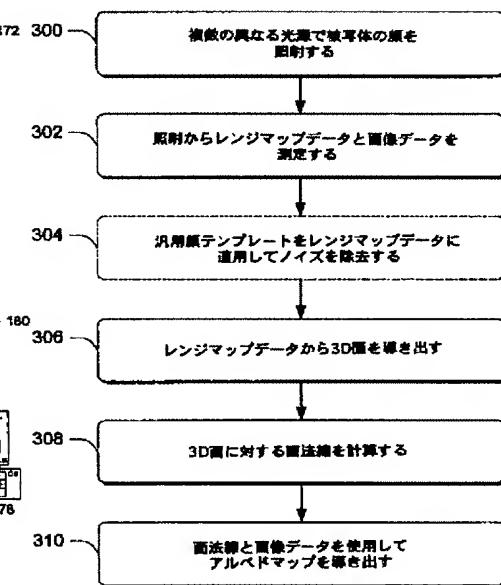
10 604 レシーバ(受信側)

606 再構築モジュール

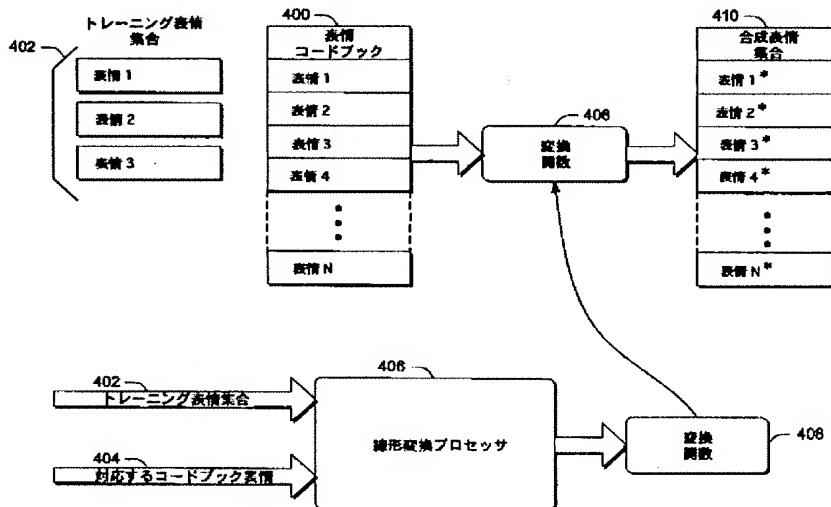
【図1】



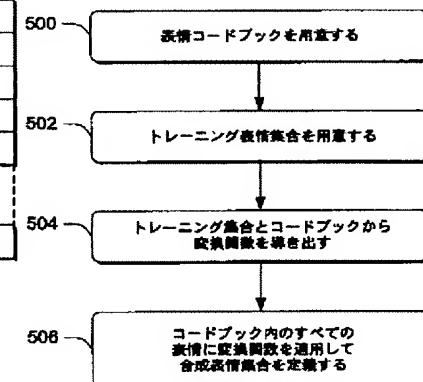
【図3】



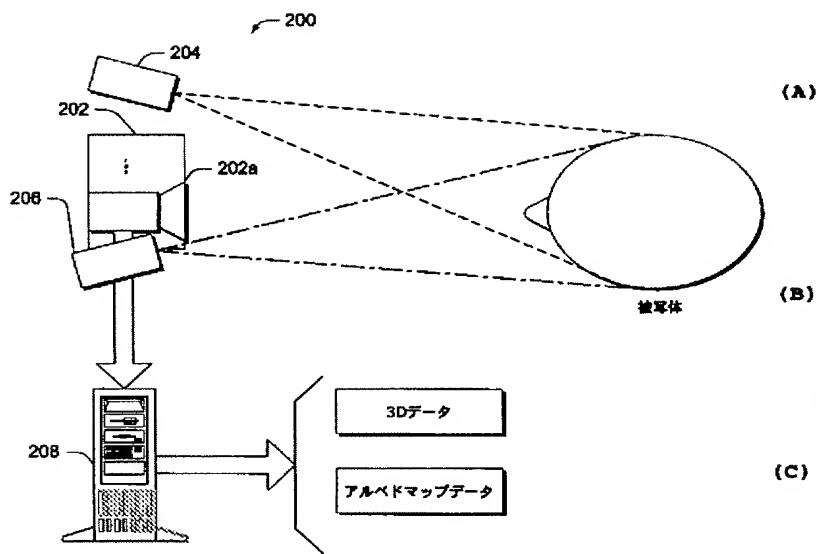
【図4】



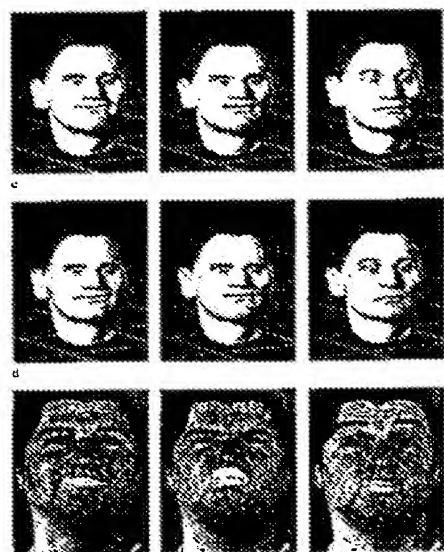
【図5】



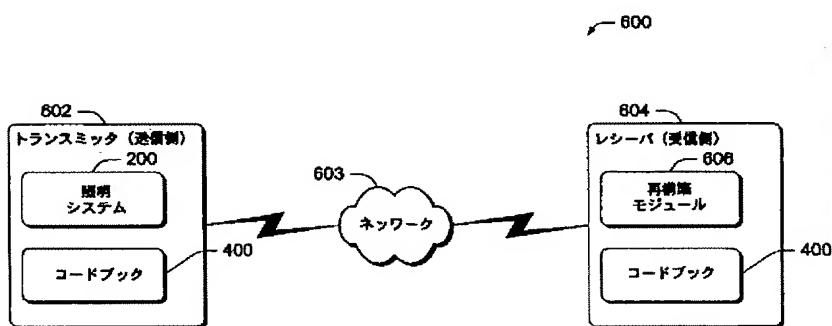
【図2】



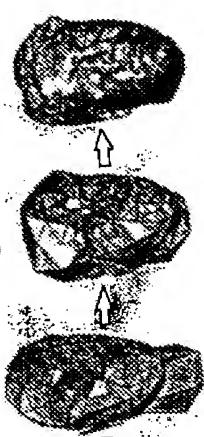
【図7】



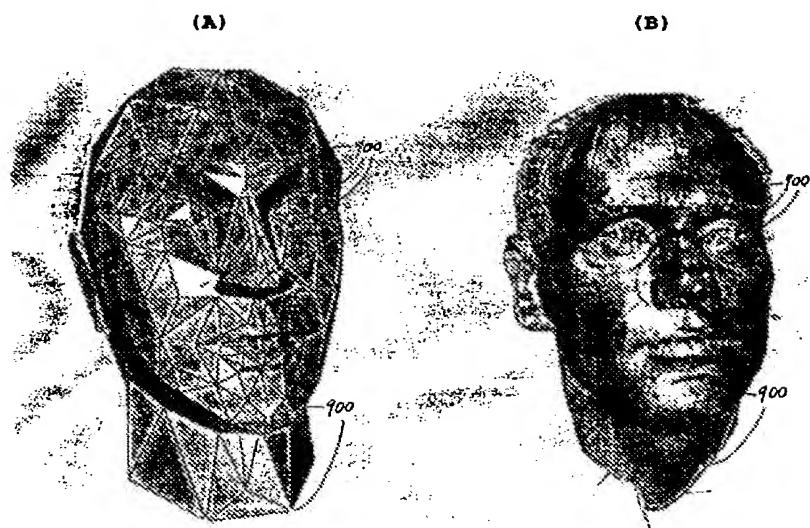
【図6】



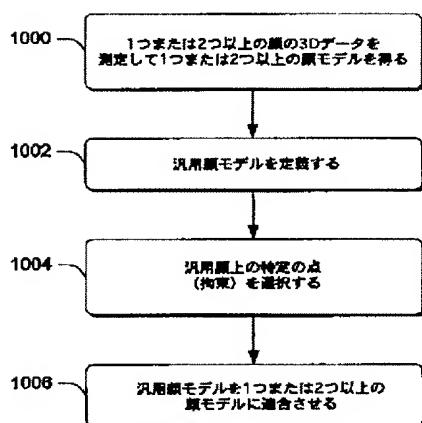
【図8】



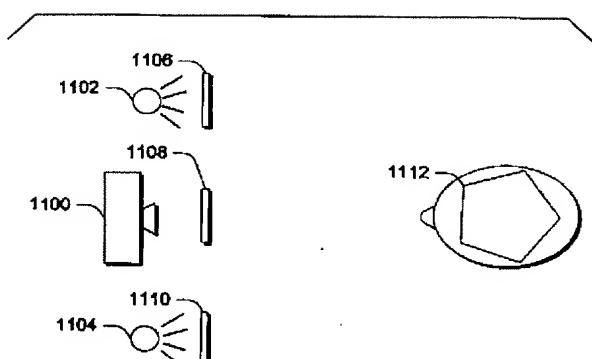
【図9】



【図10】



【図11】



【図12】

(A)



(B)



(A) (B) (C)



1行目

【図13】

(A)

(B)



(A) (B) (C)



2行目

【図18】

(A) (B) (C)



3行目

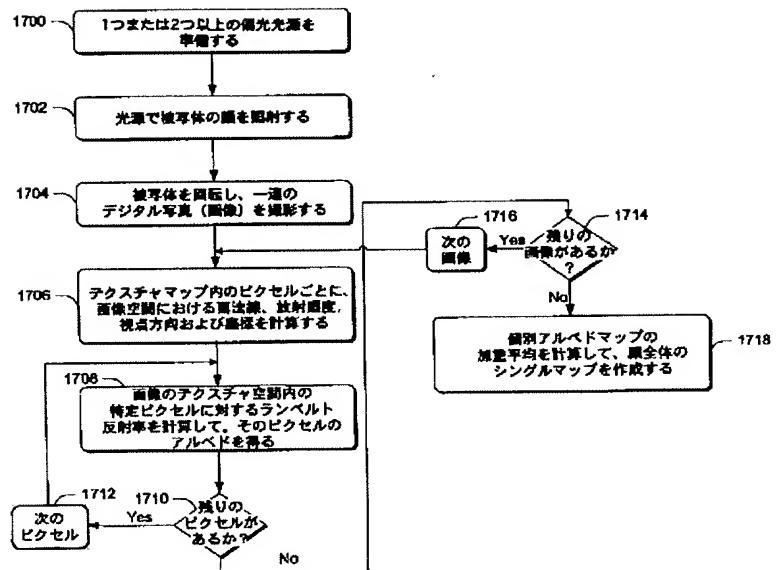
【图 1-4】



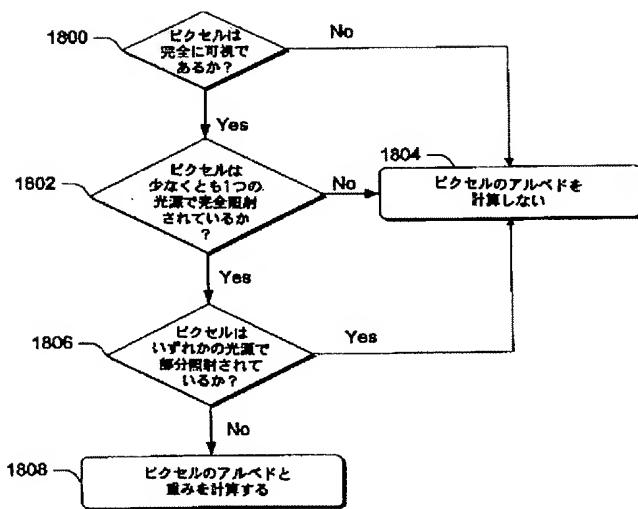
[図15]



【図19】



【図20】



フロントページの続き

(72)発明者 ブライアン ケイ. ジェンター
 アメリカ合衆国 98052 ワシントン州
 レッドモンド ノースイースト 41 スト
 リート 16725

(72)発明者 サシ ラグフパティー
 アメリカ合衆国 98052 ワシントン州
 レッドモンド 149 コート ノースイー
 スト 6110

(72)発明者 カーク オリニク
 アメリカ合衆国 98052 ワシントン州
 レッドモンド 157 コート ノースイー
 スト 13517

(72)発明者 シン ピン カン
 アメリカ合衆国 98052 ワシントン州
 レッドモンド ノースイースト 100 コ
 ート 18217

Fターム(参考) 5B050 AA08 AA09 BA08 BA09 BA12
 CA08 EA19 EA27 EA28 EA30
 FA02
 5B057 AA20 BA02 BA08 CA13 CA16
 CB13 CB16 CC01 CD11 CE08
 5L096 AA06 AA09 BA08 BA18 CA05
 CA17 DA02 FA66